Heterogeneous Chemical Kinetics and Particle Nucleation in Interstellar and Atmospheric Environments

Jayesh Bhatt

Department of Physics & Astronomy
University College London

Thesis submitted to
The University of London, UK

For the degree of
Doctor of Philosophy

2004

UMI Number: U602422

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U602422
Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Dust is believed to play a significant role in the evolution of interstellar clouds and hence in processes such as star formation. The physics involved is similar to that responsible for terrestrial aerosols.

Certain chemical reactions in interstellar conditions may only occur on the surface of a host particle and are not viable purely in the gas phase. The traditionally used rate equations approach to describe these reactions fail to account for the statistical fluctuations in the reactant populations, which would be significant in situations where the mean population may be well below unity. This can easily occur in interstellar conditions and quite often in reactions catalysed by terrestrial aerosols. This thesis considers a master equation approach that provides a stochastic description of heterogeneous chemical kinetics and demonstrates that classical kinetics may have been overestimating the reaction rates by one order of magnitude under interstellar conditions.

The same idea can be extended to study mantle growth on dust surfaces. Traditionally, this is described using a classical description of nucleation kinetics, generally suitable for large systems. Again, this can be unreliable for heterogeneous nucleation taking place on small particles under low vapour concentration where the mean population of adsorbed nucleating species could be of order unity or less. The the-

sis explores a stochastic description of heterogeneous nucleation kinetics and solves the arising equations numerically to demonstrate that the stochastic nucleation rate could be significantly different from that derived using the traditional approach.

The chemical composition of interstellar dust has for long puzzled experts. The key to determining this lies in an accurate description of the physical processes underlying the formation of these particles. Magnesium oxide is considered to be one of the major candidates as the primary nucleating material, but recently doubts have been cast over this. However, the models employed in reaching that conclusion seem to be rather inaccurate. The thesis attempts to calculate free energies of molecular clusters using newly designed potential models for MgO. It is found that MgO is probably not the primary nucleating dust species in stellar winds.

Acknowledgements

This research was carried out under the supervision of Dr. Ian J. Ford. My sincere thanks to him for his constant guidance and encouragement throughout the project, sometimes even contributing from his personal time during weekends. Thanks also to my second supervisor Dr. John H. Harding for his useful inputs during the course of the research. Special thanks to Dr. Mark Wilson of the Chemistry Department for lending me his molecular dynamics code of the CIM + PIM model and then answering my relentless queries about the use of the code. I must acknowledge the contribution of Prof. Alex Lushnikov of the Karpov Institute of Physical Chemistry, Moscow, whose initial work on heterogeneous chemical kinetics acted as a spring-board for further work regarding cosmic dust in this thesis. I have benefited greatly from inspiring discussions with Prof. David. A. Williams, one of the pioneers of astrochemistry. My friend Sascha Khakshouri deserves appreciation for critically reading parts of the thesis and suggesting ways of making things clearer.

This project was funded by the Engineering and Physical Sciences Research Council (EPSRC), to whom I am grateful. It was also kind of the UK Aerosol Society to give me the C. N. Davies Award for this work.

Finally, I must mention my family for their unconditional support during the long period of this research.

JB
London
Spring 2004

Contents

List of Figures

Glossary

The following is a list of acronyms and mathematical symbols used, except for the trivial and self-explanatory ones. The middle column lists the Chapter or the Section where the given entry is first mentioned, or in some cases, discussed prominently.

Symbol	First appearance	Meaning
$\bar{x}$		Mean value of a given quantity x.
$\langle \ldots \rangle$		Averaging a quantity over statistical fluctuations.
$\alpha, \beta, \gamma, \delta$	6.3	When appearing in index, each of these can take on any of the coordinate values x, y or z.
$\boldsymbol{\alpha}_\ell$	6.2	Dipole polarizability of the ℓ-th ion.
$a_{\ell j}$	6.3	Parameter characterising the shape of ionic charge density.
AIM	6.3	Aspherical ion model.
$A_{\ell j}$	6.2	Parameter that measures the repulsive strength of the Born-Mayer potential.
β^*	4.2	Rate of attachment of monomers to a cluster of critical size i^*.
β_i	4.1	Rate of monomer attachment to a cluster of size i.
β_i'	4.1	Reaction rate between monomers and an i-cluster. $\beta_i = \beta_i' n_1$
$\mathbf{B}$	6.3	Dipole-dipole-quadrupole hyperpolarizability.
c	3.2	Concentration of gas phase molecules surrounding a host particle.
$C_6^{\ell j}$	6.3	Dipole-dipole dispersion parameter.

Symbol	First appearance	Meaning
$C_8^{\ell j}$	6.3	Dipole-quadrupole dispersion parameter.
CIM	6.1	Compressible ion model.
$\mathbf{C}_\ell$	6.3	Quadrupole polarizability of the ℓ-th ion.
$\delta\sigma_\ell$	6.3	Change in the instantaneous radius of the ℓ-th ion .
D	6.3	Half the energy of reference ion without compression.
$\mathbf{D}'$	6.4	Mass-weighted dynamical matrix of force constants.
$\mathbf{E}_j$	6.2	Electric field at the position of the j-th ion.
$\mathcal{E}(i)$	4.2	Total energy of an i-cluster.
$F(z)$	3.3	The generating function introduced to simplify the kinetics of the heterogeneous reaction A + A $\rightarrow$ C. The first and second derivatives of this function give the mean reactant population and the reaction rate.
$F(x,y;t)$	3.3	The generating function introduced to simplify the kinetics of the heterogeneous reaction A + B $\rightarrow$ C.
$F_{i,j}$	3.4	Shorthand for the generating function value $F(x_i, y_j)$.
$\mathcal{F}(i)$	4.2	Helmholtz free energy of an i-cluster.
$\mathcal{F}_{\text{mon}}$	4.2	Helmholtz free energy of a monomer.
γ_i	4.1	Rate of monomer detachment from a cluster of size i.
$\Gamma(z)$	3.3, App. A	Gamma (factorial) function, defined for $\Re z > 0$.
$G(x,y;t)$	3.4	Equivalent to $\ln[F(x,y;t)]$; useful in obtaining a numerical solution to the kinetics of the heterogeneous reaction A + B $\rightarrow$ C.
G_m	3.4	Shorthand for $G(x,y;t_m)$, the function G at time-step t_m.
ζ	3.3	A mathematically convenient parameter, expressed as $\zeta = 2\sqrt{X(\nu+1)(1+z)}$.
h	4.2	Planck's constant (6.6262×10^{-34} J s).
$\hbar$	4.3	$h/2\pi$ (1.0546×10^{-34} J s).
$\mathcal{H}$	4.3	Hamiltonian of a system of particles.
$\boldsymbol{\theta}_j$	6.3	Induced quadrupole moment at the j-th ion position.
i	3.4	In Chapter 3, i is an index marker.

Symbol	First appearance	Meaning
	4.1	From Chapter 4 onwards, i is the number of monomers in a molecular cluster.
i^*	4.1	Critical cluster size required for nucleation to occur.
$i_{\max}$	4.1	Maximum cluster size allowed in a nucleating system.
I_1, I_2, I_3	4.3	Three principal moments of inertia of a crystal about the centre of mass.
I_{mon}	4.4	Moment of inertia of a monomer about its centre of mass.
IGM	Chap. 1	Intergalactic medium.
ISM	Chap. 1	Interstellar medium.
$I_\nu(z)$	3.3	Modified Bessel function of the first kind with the index ν being a real number and z being complex.
j	5.2	Rate of monomer attachment on particle surface.
j_A, j_B	3.2, 3.3	Rate at which A and B molecules are adsorbed on the particle surface respectively.
j_C	3.3	Production rate of the dimer C in heterogeneous chemistry.
J	4.1	Nucleation rate; the rate at which cluster of critical size i^* are created.
J_V	4.6	Nucleation rate per unit volume.
κ	3.2	Reaction rate constant such that κ/V is the reciprocal timescale of the heterogeneous reaction.
$\boldsymbol{\kappa_\ell}$	6.3	Parameter of quadrupole symmetry of the ℓ-th ion.
k_B	3.5	Boltzmann's constant (1.3807×10^{-23} J K^{-1}).
k_H	3.2	Henry's Law constant.
$K_\nu(z)$	3.3	Modified Bessel function of the second kind. The index ν is real and z can be complex.
λ	4.2	Thermal de Broglie wavelength.
	5.2	Evaporation rate of monomers from particle surface.
λ_A, λ_B	3.3	Evaporation rate of the molecules of species A and B respectively from the particle surface.
λ_k	6.4	k-th eigenvalue of the matrix $\mathbf{D}'$.

Symbol	First appearance	Meaning
LH	3.1	Langmuir-Hinshelwood.
$\boldsymbol{\mu}_j$	6.2	Induced dipole moment at the position of j-th ion.
μ_v, μ_c	4.2	Chemical potential of the vapour and the condensed phase respectively.
$\mathbf{M}$	6.4	Diagonal matrix containing the masses of oscillators in a crystal.
$M_\odot$	Chap. 1	One solar mass.
ν	3.3	Parameter that represents the adsorbed phase branching ratio in the reaction A + A $\rightarrow$ C. It is given by $\nu = (2V\lambda_A/\kappa) - 1$.
$\boldsymbol{\nu}_\ell$	6.3	Parameter of dipole symmetry of the ℓ-th ion.
n_i	4.1	Classical mean population of an i-cluster.
N	3.3	Number of reactant molecules present on a particle surface in heterogeneous reactions.
	6.2	Total number of ions in a crystal.
N_A, N_B	3.3	Number of reactant molecules of species A or B present on a particle surface.
N_i	5.2	Exact number of i-clusters present on a particle surface.
ξ	5.4	Parameter that characterises the size of the host particle in heterogeneous nucleation.
p	3.2	Partial pressure of a species in the gas phase.
$\mathbf{p}$	4.3	Momentum vector of an atom.
	4.2	Vapour pressure.
p_{sat}	4.2	Saturated vapour pressure.
$P_i(N_i)$	5.2	Probability distribution for cluster populations.
$P_{\ell i}(N_\ell, N_i)$	5.2	Joint probability distribution for two cluster populations N_ℓ and N_i.
PIM	6.1	Polarizable ion model.
Q_ℓ	6.2	Electric charge of the ℓ-th ion.
ρ_1	6.5	Initial number density of monomers when no i-clusters have been formed.

Symbol	First appearance	Meaning
ρ_i	4.6	Classical mean population density of an i-cluster.
$\rho_{\ell j}$	6.2	Steepness of the repulsive part of Born-Mayer potential.
ρ_v	4.2	Number density of monomers in the vapour phase.
ρ_{sat}	4.2	Number density of monomers in a saturated vapour.
$\mathbf{r}$	4.3	Position vector of an atom.
$r_{\ell j}$	6.2	Separation between the ℓ-th and the j-th ion.
RIM	6.3	Rigid ion model.
σ_ℓ	6.3	Radius of ℓ-th ion.
$\bar{\sigma}_\ell$	6.3	Reference value of ionic radius.
S	3.5	Sticking coefficient for molecules striking a host particle.
$\mathbf{S}^{(1)}, \mathbf{S}^{(2)}$	6.3	Interaction tensors that account for aspherical ion deformations.
$\mathcal{S}(i)$	4.2	Entropy of an i-cluster.
T	3.5	Temperature of the gas phase.
$\mathbf{T}^{(1)}$	6.3	Charge-dipole interaction tensor.
$\mathbf{T}^{(2)}$	6.3	Dipole-dipole interaction tensor.
U	6.2	Potential energy within a molecular cluster.
U_0	4.3	Potential energy of a crystal when the atoms are at their mean positions.
U_0^{mon}	4.4	Potential energy of a single monomer when the individual atoms within it are at their mean position.
v	3.2	Thermal velocity of reactant gas molecules.
V	3.2	Volume of the host particle in heterogeneous reactions.
	4.3	In nucleation, this is the volume of the 'container' in which a given cluster can translate freely.
ω_k	4.3	Angular frequency of the k-th vibrational mode of a crystal.
ω_{mon}	4.3	Angular frequency of vibration of a monomer.

Symbol	First appearance	Meaning
$W(N,t)$	3.3	Probability of finding N reactant molecules on the particle surface at time t.
$W(N_A,N_B;t)$	3.3	Probability of finding N_A and N_B molecules of species A and B respectively on the particle surface at time t.
$W(\{N_i\};t)$	5.2	Probability of finding N_1 monomers and in general N_i i-clusters on the particle surface at time t (p. 100).
$\mathcal{W}(i)$	4.2	Work of formation required for an i-cluster.
$\mathcal{W}^*$	4.2	Critical work of formation.
$\mathcal{W}_{\text{eff}}(i)$	4.6	Effective classical work of formation of an i-cluster.
x	3.3	One of the arguments of the generating function $F(x,y;t)$.
X	3.3	A parameter that represents the gas-particle equilibrium constant in the reaction $\text{A}+\text{A}\rightarrow\text{C}$. It is given by $X=j_A/\lambda_A$.
y	3.3	One of the arguments of the generating function $F(x,y;t)$.
z	3.3	Argument of the generating function $F(z)$.
$\mathcal{Z}$	4.2	Zeldovich factor.
Z	4.3	Partition function.

Chapter 1

Introduction

The Universe exhibits its material content in the form of structures ranging all the way from tiny particles of minuscule scale to gigantic stars and galaxies. The wonderful irony is that the large structures are thought to have been built as a result of painstakingly slow evolution of microscopic entities such as dust particles over millions of years. To find that a lively hot star that measures millions of kilometres in diameter would have had its origin in sub-micron dust particles wandering around in lonely, cold, dark space in the past, can be a great exercise of imagination. The driving processes within the giant structures themselves occur at microscopic and nanoscopic scales.

What is generally known as the 'empty space' is, in fact, not all that empty. There exists sparsely distributed gaseous and dusty material in the vast spaces between stars, known as the *interstellar medium* (ISM). Although sparse, this material is by no means little in amount when considered in aggregate. Crude estimates suggest that while the total mass of all stars in the Milky Way Galaxy is of the order of $10^{11} M_{\odot}$, the ISM makes up some $10^{9} M_{\odot}$ in the Galaxy, though a major proportion of the total Galactic mass ($\geq$ 90%) is attributed to the yet-mysterious 'dark

matter' in order to explain the total mass of $10^{12} M_{\odot}$ [1]. Here $M_{\odot}$ denotes one solar mass. Ordinary stars are surrounded by hot magnetised plasma that seethes out into space in thermally powered winds, feeding the material into the ISM. Within the disks of spiral galaxies, dusty clouds of molecular gas are often seen collapsing under their own gravity to form new stars. Dying stars eject enormous amounts of dusty material back into the interstellar medium, or even into the space between galaxies.

Amazing as it may all seem, it may also be bemusing as to why we should worry about chunks of worthless dust, which are so far from us that they cannot possibly have any effect on our life whatsoever. Among the many possible answers, perhaps the most captivating one – and the most publicly quoted one – is that every piece of matter in the solar system would have been a part of the interstellar medium and almost every atom in the human body would have been created in some violent stellar explosion such as a supernova. We are literally made of stardust.

Studies concerning cosmic dust began in the mid-19^{th} century as a quest to quench the curiosity about this mysterious looking isolated matter in space, although little progress was made until the beginning of the 20^{th} century. Later it became apparent that there was some rich chemistry going on in these clouds and the branch called 'astrochemistry' was born in the 1960s. Revelation of complex molecules in the interstellar environments has led to further recognition that a full development of 'astrobiology' would shed light on some fascinating facts about the origins of life itself.

Diffuse matter exists in the Universe over length scales ranging anywhere between 10^{13} m (smaller than the Solar System) and 10^{22} m (which encompasses whole clusters of galaxies) [2]. The upper value is when one considers not only the inter-

stellar, but also the *intergalactic medium* (IGM), which exists as a result of active galactic nuclei shooting away jets of hot plasma (ionised hydrogen) perpendicular to the galactic plane. However, the intergalactic material does not appear to contain any dust [3] and the plasma is so tenuous in the IGM that it almost never cools again to participate in useful chemistry due to the extremely small rate of collisions within it. A hydrogen nucleus could travel a distance of our galaxy in the IGM before encountering another of its kind. Therefore, we shall only consider the interstellar medium here since it holds great significance in relation to the process of star formation.

The dusty environment surrounding many of the giant stellar objects in the form of an atmospheric shell is quite analogous to the Earth's atmosphere, though the physical conditions may vary significantly between these two cases. For instance, in the plane of the Milky Way Galaxy, the mean density of gas is around 3×10^5 m^{-3} and inside a potential star formation region, this can be about 10^{12} m^{-3} [2]. These are negligible concentrations compared to about 2×10^{24} m^{-3} that we normally encounter in the Earth's atmosphere. Further, atomic collisions in the denser ISM occur on a timescale of several days, compared to a few nanoseconds in the terrestrial atmosphere. But there are also striking similarities that are hard to ignore. Gas and dust grains in space follow an evolutionary path in a very similar fashion to planetary atmospheres. Dust clouds in interstellar space contain a range of species from atomic gases to dust grains. These clouds may have been formed as a result of material ejection from stellar outflows. In the vicinity of cool atmospheres of red giants and supergiants, we find perfect conditions for the nucleation and growth of dust grains to occur. The grains may grow as a result of gas phase atoms and molecules adsorbing permanently on the grain surface. This is known as the

mantle growth of dust grains. Another mechanism for the grain growth is through *coagulation*, whereby the dust particles themselves may collide and occasionally stick together, growing from sub-micrometre sized grains to small pebbles. If the interstellar conditions are right, these pebbles may grow and eventually collect into even macroscopic objects, whereupon gravity assists the building process of giant structures like planets and stars. The possibilities are far more diverse in outer space than in terrestrial atmospheres and the densities of the material involved may be much lower in space with no gravity. However the underlying physics of grain condensation, nucleation and growth from microscopic size is very much the same in both cases.

It has been generally recognised that small dust particles participate in several chemical reactions, which would otherwise be impossible to come about in the absence of these particles. Certain chemical reactions can only occur on the surface of suspended particulate matter, rather than purely in the gas phase, in order to ensure the conservation of energy and momentum of the reactant molecules. In such a case, the dust acts as a 'catalyst', such as in the case of the simple reaction $H + H \rightarrow H_2$. Similarly, atmospheric aerosols assist important aqueous reactions like $HSO_3^- + OH \rightarrow SO_3^- + H_2O$ and there are many more examples of heterogeneous chemical reactions in our atmosphere [4]. In fact, some aerosol scientists often like to term cosmic dust as simply 'cosmosols'.

The original stimulus behind this thesis was the formation and the evolutionary role of dust in cosmic environments. However, given the similarities between the interstellar and atmospheric dusty environments, the theoretical ideas developed here have been applied to both these cases. Any knowledge obtained from one perspective may conveniently be applied to the other one, so while modelling cosmic

dust, some of the known features of atmospheric aerosol science have also been instrumental.

The issue mentioned above can be categorised into mainly three aspects: (i) formation of cosmic dust from vapour phase through homogeneous nucleation, (ii) growth of dust grains by means of heterogeneous nucleation of mantles and (iii) recombination chemistry occurring on the surface of dust particles. Each of these aspects is a comprehensive problem in itself and yet they all combine together in order to give a more complete picture of the subject. This thesis will touch upon each of these in one way or another and attempt to fill some of the gaps in our knowledge.

Since the cosmic dust problem is intrinsically related to the physics of star formation, in the next Chapter we begin by reviewing the accepted model for the cycle of star formation from interstellar dust and its eventual dispersion back into the ISM. In Chapter 3 we develop a model for grain surface chemical reactions that is suitable for the low density conditions of the ISM. A major part of this work has been published in Ref. [5], which was co-written by the author. Chapter 4 notes some of the key ideas of the theory of nucleation, which will be useful in the subsequent two Chapters. The model developed in Chapter 3 is further extended in Chapter 5 to describe heterogeneous mantle growth on dust grains in conditions where the gas and dust concentrations may be very small. This has been published in Ref. [6]. The question of which chemical species nucleates first to form dust in space has been addressed in Chapter 6, where we examine the possibility of MgO as a candidate by calculating the free energy of formation of MgO clusters in circumstellar shells. Finally, the major conclusions are listed in Chapter 7.

Chapter 2

Star Formation and Evolution

Cosmic dust is seen by the observational astronomer as an irritating factor that obscures vision and hampers efforts to obtain trustworthy spectra of cosmic objects. For the astrophysicist, however, dust is a crucial ingredient of the cosmic soup, for it is dust that connects two seemingly separate generations of stars through a rich network of reactions.

Stars are formed from the huge amount of gaseous and dusty material available in the form of the interstellar medium and when a star reaches the end of its life-cycle, the stellar material is returned back into the same ISM. Thus the cycle of construction and destruction of a star somewhat resembles the classic chicken-and-egg problem; one could not have come into existence without the other one having existed previously. How exactly the first stars, or the first interstellar medium, appeared in the universe is a complicated problem for cosmologists [7] and is well beyond the scope of this brief overview. For us, it is sufficient to focus on a picture where the existence of both stars and the ISM is sustained through a mutual exchange of material between the two.

2.1 The main sequence

During its lifetime, a star's stability is guaranteed due to a balance between its self-gravity and the outward force due to the energy released in burning its nuclear fuel in the core. Hydrostatic equilibrium is maintained due to the fact that as we approach the centre of the star, the gravitational pressure increases, but so does the outward pressure originating from the heat of nuclear burning, hence preserving the stability [3]. The star is most stable when hydrogen within it is being burnt to produce helium through nuclear fusion via the proton-proton chain (which is the predominant process in the Sun) and/or the CNO cycle (if the star has carbon in it). During this phase, the star is said to be on the *main sequence* of its lifecycle, which is the longest stage in its life. The time a star spends on the main sequence strongly depends on its mass. The Sun has a main sequence lifetime of 10 billion years (about 80% of its total lifetime), of which 5 billion years have passed. In contrast, a star with $15M_{\odot}$ will stay on the main sequence for only about 10 million years [8]. Also, the Sun at present has a core temperature of around 15 million K, whereas its surface temperature is some 5,700 K. Compared to this, the higher mass O-stars have surface temperatures in the range of 28,000 – 50,000 K during the main sequence.

2.2 Star death and dispersion

When all the hydrogen is burnt up in the core, the star moves off the main sequence. Adjustments are needed in order for the star to survive further and this leads to fundamental changes in its structure. The inert core contracts rapidly, which in turn releases gravitational energy, heating the surrounding layers of hydrogen to the

point where hydrogen burning recommences not in the core, but in a shell around the core. The new outpour of energy pushes the outer layer of the star further and further. The expanding shell eventually cools, which means the star appears reddened in addition to having attained a much larger size. In other words, the star becomes a *red giant* [9].

The newly created helium core continues to contract until it reaches a temperature of $\sim 10^7$ K again, so that it is hot enough for the fusion of helium into carbon and oxygen to begin.

Eventually, all the helium in the core is also consumed. What happens subsequently depends on the mass of the star. In the more massive stars, contraction of the core after each fuel has been exhausted raises the temperature sufficiently to ignite a new, heavier fuel. Ultimately, a situation can be reached in which the central core has been converted to iron, while around the core, in a series of shells, increasingly lighter elements like silicon, oxygen, carbon, helium and hydrogen are being burnt simultaneously. By this time, a star of $10M_{\odot}$ would have taken the form of a *red supergiant*, with a diameter of about 100 to 500 times that of the Sun [3]. Once a star has developed an iron core of about one solar mass, no new reactions are possible. Iron has the maximum nuclear binding energy per nucleon ($\sim$ 8.8 MeV), hence it will be favoured over the slightly heavier elements that hold their nuclei less tightly [10]. At this stage, the stellar core contracts until it implodes catastrophically (in less than one second), quickly setting off a supernova explosion, the most violent event known in the Universe after the Big Bang. Copious material containing a rich variety of chemical elements is thrown off into the surrounding space. A supernova is also believed to be the production house for elements heavier

Figure 2.1: The Dumbbell nebula was the first planetary nebula ever discovered, which was in 1764, and remains the most impressive object of its kind in the sky. It is about 125,000 lightyears away. Estimate suggests that at present we are observing the shell ejection 3,000 – 4,000 years after it started. The central star is likely to be an extremely hot O-star with a surface temperature of around 85,000 K. © J. M. Anderson (www.astro.virginia.edu/~jma2u).

than iron*. The naked compact, heavy core left behind following a supernova often succumbs to its own gravity, reaching a dead end in the form of a neutron star, the most compact object ever observed directly.

*Nuclei of heavier elements are formed due to successive captures of neutrons by an existing nucleus. A captured neutron, however, will tend to β-decay into a proton and as a result, certain nuclei tend to be more stable than the others, especially in the heavier elements region. As we move towards heavier elements in the periodic table, the number of neutrons in a nucleus increasingly exceeds the number of protons. Hence for a nucleus to grow from one stable size to the next stable size, the neutron capture rate needs to be more rapid than the β-decay rate so that the right balance between the number of protons and neutrons can be struck in the next stable nucleus. As one tries to build heavier and heavier nuclei, an ever-faster feeding rate of neutrons is required. Even the hottest of the stars do not have a neutron producing mechanism that is fast enough to create elements heavier than iron. Such conditions may only be created in a supernova, in which large amount of neutrons are believed to be created in a short time span. (See, for example, Refs. [11, 12, 13])

In the lower mass stars, however, the end is less extravagant and there is possibility of a more enduring future. In stars like the Sun, the central temperature never gets high enough to progress beyond the burning of hydrogen and helium in the core and the outer shell. The contracting core gives rise to an increasing temperature, which leads to an acceleration in the shell-helium burning rates. These instabilities make the star pulsate until the outer layers of the star are separated from the core. An expanding shell of gas, known as a *planetary nebula*, is formed, which gradually disperses into space. Even the high mass stars have been observed to undergo rapid mass loss towards the end of their lifetime, forming planetary nebulae (Figure 2.1). In fact, what drives the huge mass loss during the final stages of a star's lifecycle still remains an intriguing topic of research, but what is known is that significant mass is probably lost from most stars through *stellar winds* during the later stage of evolution. The remaining core of the lower mass star cools and shrinks until it is about the size of the Earth. Because of the high core density, the matter becomes degenerate there and a *white dwarf* is formed. White dwarfs continue to cool further, since there is no internal source of energy left.

The gaseous and dusty material expelled by dying stars travels far out into space and as a result the interstellar medium is full of sparsely distributed matter. The Sun at present loses mass at a rate of some $10^{-14} M_{\odot}$ yr^{-1} due to the solar wind. Red giants and supergiants are known to throw away envelopes at rates of 10^{-7} to $10^{-6} M_{\odot}$ yr^{-1} [3, 14]. Still, the densities associated with the ISM are extremely small: on the average, one dust particle can be found in 10^6 m^3 of interstellar space [15], whereas the density of atomic hydrogen, the most abundant element in the Universe, can range anywhere between 10^3 m^{-3} (supernova remnants) and 10^{11} m^{-3} (molecular clouds) [1]. Nevertheless, the collective effect of the vast regions

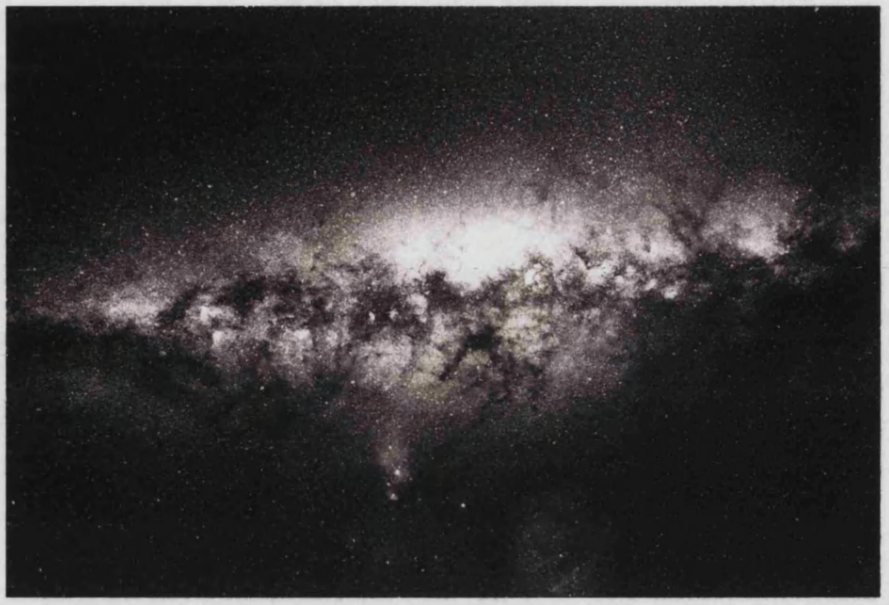

Figure 2.2: The Milky Way as seen from the Earth. The dark patches within the galactic disc are created due to dusty regions blocking the light from stars behind them. © IG Astrofotografie Bochum (aida.astroinfo.org/iab).

of interstellar dust is such that they can obstruct light from stars located behind them and create huge dark patches in the night sky (Figure 2.2). This dust can also scatter light back into space, giving rise to beautiful *reflection nebulae*. The exact nature of the scattering depends on the size, shape and composition of the dust grains. Scattered starlight makes up about a quarter of the light we see from the Milky Way, so meaningful observation of an awful lot of objects requires a good understanding of cosmic dust.

2.3 Birth of a new stellar generation

When a certain region of the diffuse ISM experiences a significant external influence, such as a shock wave originating from some colliding or exploding cosmic objects in the neighbourhood [2, 16], gravitational instabilities are set up and it collapses to

form a *molecular cloud.* It is noteworthy that very massive stars, which might appear to be selfishly short-lived and extremely violent, in fact do a good job of stirring and mixing the interstellar clouds and pave the way for future generations of star systems. Molecular clouds can have diameters anywhere from 40 to 350 lightyears (Figure 2.3). Gravitational collapse of a cloud can occur at several different places in the interstellar region. As the molecular cloud shrinks, molecules within it are heated and the dust grains radiate at infrared wavelengths. As long as this radiation can escape into the surrounding space, the cloud remains cool, the pressure stays low and the collapse continues as gravitational free fall. Such a cloud is known as a *protostar.*

Figure 2.3: N70 in the Large Magellanic Cloud: A luminous bubble of interstellar gas, measuring about 300 light-years in diameter. It was created by winds from hot, massive stars and supernova explosions and the interior is filled with tenuous, hot expanding gas. An object like N70 provides astronomers with an excellent opportunity to explore the connection between the lifecycles of stars and the evolution of galaxies. © European Southern Observatory (www.eso.org).

Figure 2.4: The largest emission nebula in the sky, the Tarantula Nebula (also known as NGC 2070 or 30 Doradus) is located in the Large Magellanic Cloud (LMC), one of the satellite galaxies to our own Milky Way system. Seen far down in the southern sky at a distance of about 170,000 light-years, this beautiful nebula measures more than 1000 light-years across. Thousands of stars are being born in this fertile region of gas and dust. © European Southern Observatory (www.eso.org).

At some stage, however, the cloud becomes optically thick and the radiation starts getting trapped, which not only heats up the cloud, but also slows down the collapse dramatically as hydrostatic equilibrium is established. Thereafter, the star is treated as a pre-main-sequence (PMS) star. The temperature is still so low that its opacity is relatively high. Hence, convection rather than radiation is the means for transporting the energy outwards. A newly formed PMS star is fully convective from centre to surface and the effective transport of energy makes it extremely luminous. This is the reason why star forming regions, such as the spectacular 30 Doradus, are often so colourfully bright (Figure 2.4).

Eventually, the opacity drops to a level where radiative transport of energy

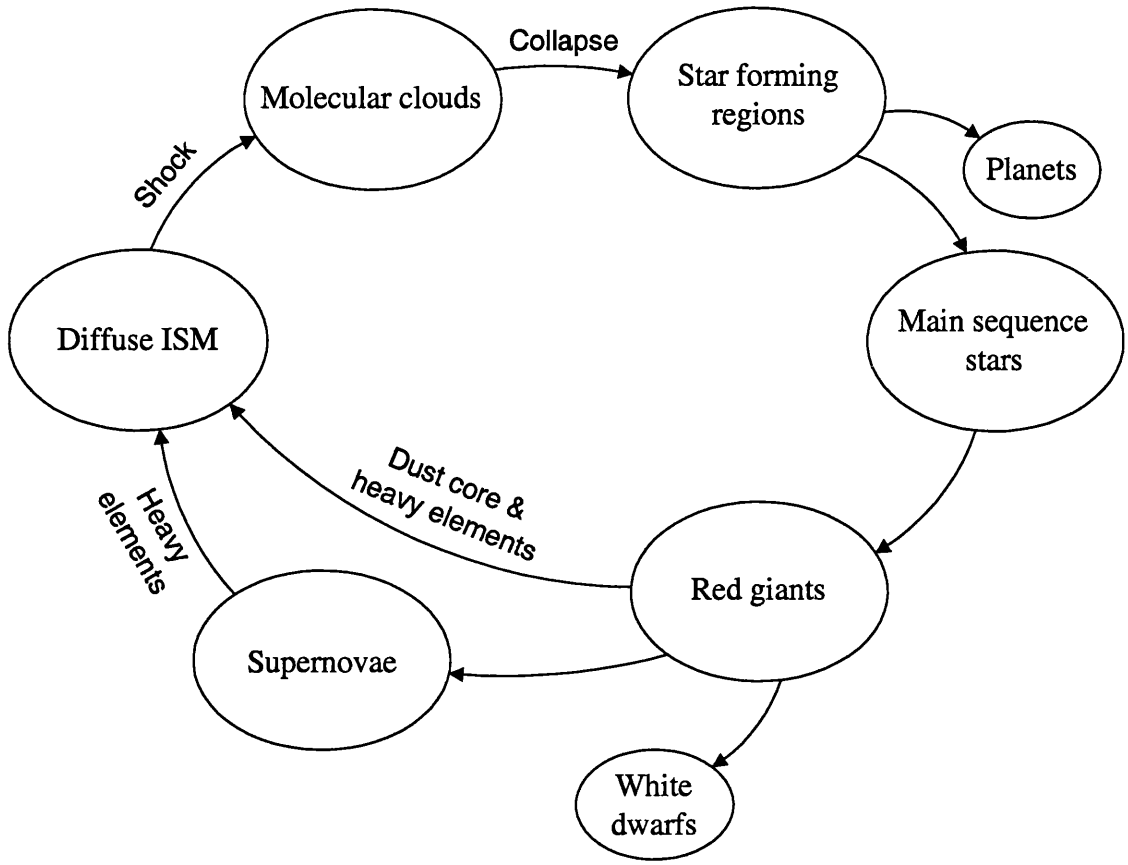

Figure 2.5: An illustration of the evolutionary cycle of star formation and destruction from interstellar dust. Based on Ref. [17].

dominates over convection. This happens first at the core and gradually spreads outwards. When the core heats up to a few million Kelvin, thermonuclear reactions trigger off and a full-fledged star is born. It is believed that most probably a high mass star is first formed close to the front edge of the molecular cloud where the shock was incident. This in turn can send further shocks back into the molecular cloud, which then act to trigger further formation of smaller star systems. If any isolated globules are created containing mass that is insufficient to form stars, they gradually become dispersed in the interstellar medium due to actions of stellar winds and shocks. A schematic representation of the entire cycle of star formation and death is shown in Figure 2.5.

It must be stressed that even after more than a century of research, the exact mechanism driving star formation remains mystifying. Nevertheless, a few facts are known for certain. Star formation is not only an extremely slow process, but also very inefficient. A parent molecular cloud can be as massive as $10^5 - 10^6 M_\odot$, whereas the mass of each star formed would be typically in the range $0.1 - 50 M_\odot$. This means that only a small fraction of the material is consumed from a molecular cloud after each star birth. This explains the massive amount of interstellar medium we are able to observe at any given epoch. It also highlights long-lasting possibilities of further star formation and the need to have a good understanding of the dust that plays a crucial role in it.

Chapter 3

Stochastic Kinetics of Grain Surface Chemistry

3.1 Heterogeneous chemical reactions

Particulate matter plays a key role in the chemical evolution of its gaseous environment [4, 18]. Some reactions can only occur in the presence of a third body rather than purely in the gas phase (for example, see Ref. [19]). This is familiar in atmospheric heterogeneous chemistry and in astrochemistry too. An example is the simple reaction $H + H \rightarrow H_2$, an important reaction in the evolution of the cosmic chemical environment. It was pointed out by Gould & Salpeter [20] in 1963 that gas phase reactions cannot be an efficient enough process for the production of H_2 under interstellar conditions. The H atoms cannot recombine directly by two-body encounters in the gas phase because there is no feasible method of losing binding energy. Three-body encounters would be extremely unlikely, therefore Gould & Salpeter advocated the idea that dust grains may act as the third body catalysts for such reactions. The surface of a grain provides both a substrate on which atoms

can get trapped and a heat reservoir that absorbs excess energy.

For simplicity, we shall use the term *molecule* for both atoms and real molecules since for the present discussion the exact identity and constitution of the reactant species are immaterial. In the same context, we shall refer to a dust particle simply as a *particle.*

Primarily there are two physical mechanisms that could lead to chemical processes inside or on the surface of aerosol and interstellar particles. In the first one, a gas phase molecule incident upon the grain surface lands directly on top of another adsorbed molecule and the two may react in order to produce a dimer. This is the so-called Eley-Rideal mechanism and Farebrother *et al.* [21] have investigated the formation of molecular hydrogen under this scheme.

The study presented here, however, is concerned with a second mechanism, known as the Langmuir-Hinshelwood (LH) mechanism. Here the particle is considered to be a tiny reaction vessel within which chemistry takes place. A dust particle surrounded by gas phase molecules will have a certain probability of adsorbing such a molecule onto its surface over a given period of time. The adsorbed molecule may wander on the surface of the particle (or within its volume), during which it may encounter another molecule that happened to be there from a previous capture. In such a case, the two molecules may combine to produce a reaction product, which might then desorb. Again for simplicity, we shall refer to the reactants being able to move on a particle surface alone, rather than within it.

As an attempt to study grain surface reactions via the LH mechanism, mainly two theoretical approaches have been developed. The first of these is that of Allen & Robinson [22], according to which the surface chemistry is "accretion limited",

i.e. the chemistry is limited by the adsorption rate of reactant molecules onto the grain surface. If two reactive molecules are present on the surface and if at least one of them is able to move around the surface quickly enough, then the two can combine *before a third molecule lands on the surface.* Provided that this condition is fulfilled, the diffusion rate holds little significance and the chemistry can be characterised solely by the adsorption of gas phase molecules. This approach is particularly promising for the low density conditions of the interstellar medium. It can be, and has been, implemented with the aid of Monte Carlo simulations [23, 24, 25, 26]. However, these attempts have only resulted in time-independent chemical models and the desired time-dependency has not been achieved so far under this approach. Furthermore, it is limited in scope due to the basic assumption of accretion limitation.

In the second approach to the LH mechanism, the ideas of standard gas phase chemical kinetics are applied. This approach was formulated by Pickles & Williams [27] and has been the more popularly used method due to its flexible nature [28, 29, 30, 31, 32]. Reactions are described by the classical rate equations of chemical kinetics, written in terms of the average number of reactant molecules present on the particle surface which is time-dependent in general. The key calculation in determining the reaction rate is the probability of a collision between reactants. This is normally obtained by multiplying the molecular flux (proportional to the mean molecular concentration) by the collision cross section (also proportional to the concentration), yielding a reaction rate proportional to the square of the mean population.

This is all very well for macroscopic reaction vessels, but as A. G. G. M. Tielens pointed out at a conference in 1995 (unpublished), when the particles have sub-

micrometre dimensions there is cause to question the rate equation approach of Pickles & Williams. If the mean number of molecules of a reactant on the particle is around unity, then there are often no attached molecules present and only on rare occasions, more than one. In other words, fluctuations in the populations will be significant. Multiplying two average populations together is unlikely to provide the correct reaction rate and the classical approach will no longer apply.

3.2 The stochastic limit

Concern about the validity of the classical approach is the greatest for reactions involving ultra-low concentration gases adsorbing onto ultrafine particles, since the adsorbed reactant populations are the smallest for these conditions. The need for a stochastic approach in the light of fluctuations in populations has long been recognised and has been addressed in different ways and contexts. Gillespie [33, 34] developed stable numerical methods for simulating the time evolution of stochastic chemical systems involving several coupled reactions. A stochastic master equation approach was applied to the chemistry occurring in micelles by Hatlee & Kozak [35, 36, 37]. Later, the possible effects of small populations on chemistry taking place on atmospheric aerosol particles was highlighted [38] and a Monte Carlo simulation based on stochastic arguments has also been constructed before [39], but a rigorous analytic insight into the kinetics of these reactions remained at large. Attempts were even made to address this problem by modifying the classical rate equations in such a way as to match the results of, for example, Monte Carlo simulations without identifying the underlying physical processes at low concentrations [40, 41, 42].

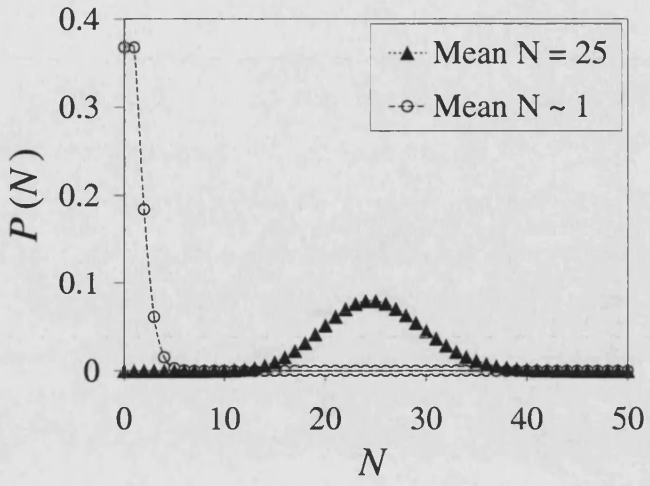

Figure 3.1: Poisson distribution for two different values of mean population ($\bar{N}$) of reactant molecules according to $P(N) = \mathrm{e}^{-\bar{N}}\bar{N}^N/N!$. When $\bar{N}$ is reasonably large (for instance, $\bar{N} = 25$), the distribution is symmetric. As $\bar{N}$ tends to unity, the plot becomes asymmetric since it is meaningless for N to take negative values. Only integer values of N are physical; the dotted lines are fitted to demonstrate the qualitative behaviour.

A more appropriate description of the kinetics would be to focus on the probability distribution for the number of adsorbed molecules, N, on the particle. In Figure 3.1, a Poisson distribution for this number is shown, in situations where the mean is large and where it is small. When the probability of 'success' of an event is large, such as in the limit of reasonably moderate or high density clouds of gas, a Gaussian shape may well approximate the actual probability distribution. However, when the success probability is small within a large sample of trials, the statistics is represented by the Poisson distribution, whose mean $\bar{N}$ is equal to the variance $\overline{N^2} - \bar{N}^2$. In low density clouds, the adsorption probability of a molecule is small and the probability of finding two reactant molecules simultaneously on the surface is even smaller. The probability of reaction is therefore proportional to $\overline{N(N-1)}$ (for $N \geq 1$), which in low density conditions may be significantly different from the estimation $\bar{N}^2$ that is normally used in the classical approach. For $N = 0$, the

reaction probability would of course be zero regardless of the way one treats the kinetics. The identity $\overline{N(N-1)} = \bar{N}^2$ always holds true for a Poisson distribution, so strictly speaking, the classical rate law is only valid if one considers that distribution. For a better accuracy of the kinetics, we ought to determine through a stochastic method how the entire probability distribution, rather than only the mean population, evolves in time.

Incidentally, it is worth asking whether populations of reactant molecules of order unity are realistically to be found in practice. Let us consider the simplest possible second order chemical process $\mathrm{A} + \mathrm{A} \rightarrow \mathrm{C}$ and a dust particle of radius a, surrounded by gas phase molecules A with a concentration c, so that there is a flux j_A of A-molecules onto the particle surface per unit time. This flux is estimated as $j_A \sim \pi a^2 v c$, where v is the thermal velocity of reactant molecules. To quantify the dimer production, we may regard the reaction as taking place in a vessel of volume $V \sim \frac{4}{3}\pi a^3$ and assign an effective reaction rate constant κ such that the reciprocal time of reaction is κ/V. Such an effective reaction rate could be used to characterise surface reactions too.

The ratio of the adsorption rate to the reaction rate is of the order of unity at the concentrations $c \sim \kappa/\pi a^2 v V$. Let us now consider particles with radius $a = 0.1\ \mu\mathrm{m}$ and a reaction coefficient of $\kappa = 10^{-20}\ \mathrm{m^3\ s^{-1}}$ [4]. We also assume $v \sim 10^3\ \mathrm{m\ s^{-1}}$. The gas phase molecular concentrations at which the ratio of rates is unity are not incredibly small: $c \sim 10^{14}\ \mathrm{m^{-3}}$, so may well be encountered in real situations. These concentrations are a few parts per 10^{11} in the terrestrial atmosphere. Concentrations of trace atmospheric constituents can fall below these thresholds and interstellar gas species are easily dilute enough. Furthermore, if the particle is smaller, or if the sticking coefficient is less than unity, then the respective threshold concentrations

are higher. This may well lead to the scenario where the population of the reactants is of order unity.

More precisely, the number of molecules N of a trace atmospheric gas dissolved in a droplet will depend on its partial pressure p, the size of the droplet V and the Henry's Law constant k_H according to

$$N = N_A 10^3 k_H p V, \tag{3.1}$$

where N_A is Avogadro's number, p is in atm, V is in m^3 and k_H is in units of $\mathrm{M\,atm^{-1}}$. Taking ozone as an example, present in the lower atmosphere typically at a concentration of 10 ppb ($p \approx 10^{-8}$ atm), with $k_H = 1.1 \times 10^{-2}\ \mathrm{M\,atm^{-1}}$ for dissolution in water, we find that $N = 0.28$ for a droplet of radius 1 μm. It would appear that any aqueous phase ozone chemistry taking place in droplets of this size is operating in a regime where the (mean) reactant population is less than unity.

The issue of heterogeneous chemistry in this limit is addressed in this Chapter with a new formulation of reaction kinetics (the so-called stochastic chemistry) that was developed independently by Lushnikov *et al.* [43, 5] and by Green *et al.* [44] as a replacement of the standard chemical kinetics approach. We begin with a simple reaction involving two molecules of the same type reacting to produce some new molecule. We shall later extend this idea to the slightly more realistic reaction between two different species. To compare the stochastic approach with the results of traditional chemical kinetics, we shall apply both of them to examples of reactions, each in interstellar medium and in terrestrial atmosphere. This will allow us to draw conclusions on the applicability of standard rate equations for describing heterogeneous chemistry in tiny reaction vessels.

3.3 Stochastic treatment of heterogeneous chemical kinetics

3.3.1 Reaction A + A → C

For reactions of the type in which two molecules (or atoms) of the same species A combine to form a product C on the surface of a spherical particle of volume V, each state of the particle can be characterised by the number N of A-molecules attached to it. We assume that there is a flux j_A of A-molecules adsorbing on the particle surface per unit time (Figure 3.2). To quantify the dimer production, we may assign a reaction rate constant κ such that the reciprocal time of reaction is κ/V. On the other hand, many of the adsorbed molecules will evaporate from the surface before they get a chance to react with another molecule, hence introducing an evaporation rate λ_A which has the dimension of inverse time.

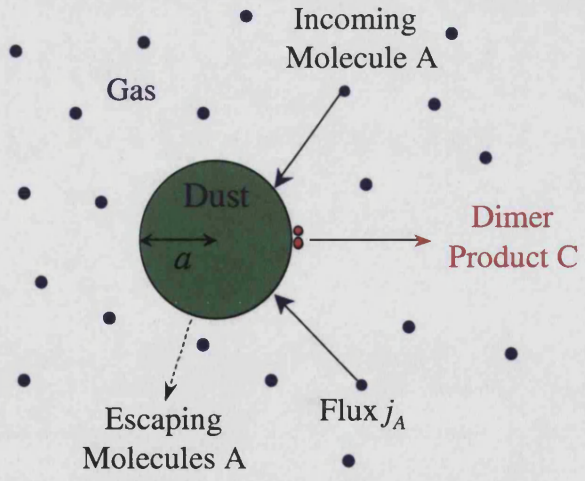

Figure 3.2: A schematic representation of grain surface reactions. A comparatively large dust particle of radius a is surrounded by a gas comprising A-molecules. The mean population of the A-molecules on the dust particle surface in equilibrium will be $\bar{N}$.

Let $W(N,t)$ be the probability of finding exactly N molecules on the particle. We can then write down a master equation governing the time evolution of $W(N,t)$:

$$\begin{aligned}\frac{\mathrm{d}W(N,t)}{\mathrm{d}t} &= j_A[W(N-1,t) - W(N,t)] \\ &\quad + \frac{\kappa}{2V}[(N+2)(N+1)W(N+2,t) - N(N-1)W(N,t)] \\ &\quad + \lambda_A[(N+1)W(N+1,t) - NW(N,t)]. \end{aligned} \tag{3.2}$$

The meaning of the terms on the right hand side of this equation is as follows. The first term, $j_A W(N-1,t)$, represents the increase in the probability of finding N adsorbed molecules through the attachment of molecules at a rate j_A to a particle containing $N-1$ molecules. The term $j_A W(N,t)$ describes the decrease in probability of finding N molecules in the droplet, through a similar attachment process to a droplet already containing N molecules. The first group of terms in squared brackets therefore describe jumps in population $N-1 \longrightarrow N \longrightarrow N+1$ due to the adsorption of a molecule from the outside.

The second group of terms corresponds to the transitions $N+2 \longrightarrow N \longrightarrow N-2$ due to the binary reaction (with the removal of product from the surface). The probability of finding $N+2$ molecules is $W(N+2,t)$ and the probability of finding N molecules increases at a collision rate κ/V multiplied by the number of pairs of A molecules that can be found amongst the $N+2$ molecules, namely $(N+2)(N+1)/2$. The third group describes the jumps $N+1 \longrightarrow N \longrightarrow N-1$ due to the evaporation of a molecule from the particle. Factors of N etc. appear where necessary to give the correct coefficients. The initial and normalisation conditions completely define the solution of Equation (3.2).

This scheme replaces the much simpler equation of classical chemical kinetics:

$$\frac{\mathrm{d}\bar{N}}{\mathrm{d}t} = j_A - \frac{\kappa}{V}\bar{N}^2 - \lambda_A \bar{N} \tag{3.3}$$

where $\bar{N}$ is the average number of A-molecules on the surface. The terms on the right hand side correspond, respectively, to adsorption, reaction and desorption. This equation can in fact be derived from Equation (3.2), by multiplying both sides by N, summing over N and regarding $\bar{N}$ as large.

3.3.1.1 Simplification of master equations through generating function

In order to simplify Eq. (3.2) let us introduce a generating function

$$F(z) = \sum_{N=0}^{\infty} z^N W(N). \tag{3.4}$$

It can be shown that in steady state (i.e. when $\mathrm{d}W/\mathrm{d}t = 0$), the master equation reduces to

$$\frac{\kappa}{2V}(z+1)\frac{\mathrm{d}^2F}{\mathrm{d}z^2} + \lambda_A \frac{\mathrm{d}F}{\mathrm{d}z} - j_A F = 0\,. \tag{3.5}$$

This can be demonstrated as follows. The first and second derivatives of $F(z)$ in Eq. (3.4) will give rise to terms in z^{N-1} and z^{N-2} within the series. If we then substitute the expressions for $F(z)$ and its derivatives in Eq. (3.5) and expand the summations explicitly, we are left with an infinite series with ascending powers of z. One may focus on some particular value of N (say $N = 4$), collect terms in z^N from this expanded series and equate it with zero, as required by Eq. (3.5). This gives

$$\begin{aligned} 0 \;=\; & \frac{\kappa}{2V}[(N+1)NW(N+1) + (N+2)(N+1)W(N+2)] \\ & + \lambda_A(N+1)W(N+1) - j_A W(N). \end{aligned} \tag{3.6}$$

Performing a similar procedure for terms in z^{N-1} gives

$$0 = \frac{\kappa}{2V}[N(N-1)W(N) + (N+1)NW(N+1)] + \lambda_A NW(N) - j_A W(N-1). \tag{3.7}$$

By subtracting Eq. (3.7) from Eq. (3.6), the steady state version of the master equation (3.2) is recovered (i.e. with $\mathrm{d}W/\mathrm{d}t = 0$). Solving the ordinary differential

equation (3.5) is therefore equivalent to solving the original master equation, but the problem is simplified considerably. This is the purpose of introducing the generating function.

3.3.1.2 Solution for the generating function

Before proceeding further, let us define two parameters X and ν as

$$X = j_A/\lambda_A \tag{3.8}$$

and

$$\nu = \frac{2V\lambda_A}{\kappa} - 1, \tag{3.9}$$

which represent the gas-particle equilibrium constant and the adsorbed phase branching ratio respectively, reflecting the gain and loss mechanisms of A-molecules. These will be convenient for later analysis. It is also useful to define another parameter ζ such that

$$\zeta = 2\sqrt{X(\nu+1)(1+z)}. \tag{3.10}$$

The solution for $F(z)$ may now be obtained through use of the following ansatz [45, p. 985-986]. We start with a trial solution of the form

$$F(z) = F_0(1+z)^{-\nu/2}\phi(\zeta), \tag{3.11}$$

where F_0 is a constant and ϕ is a yet to be determined function. By suitable manipulations, it may be shown that Equation (3.5) reduces to the following equation for ϕ:

$$\zeta^2\frac{\mathrm{d}^2\phi}{\mathrm{d}\zeta^2} + \zeta\frac{\mathrm{d}\phi}{\mathrm{d}\zeta} - (\zeta^2+\nu^2)\phi = 0. \tag{3.12}$$

The solutions to Equation (3.12) are the modified Bessel functions $I_\nu(\zeta)$ and

$K_\nu(\zeta)$ (see Appendix A). The general solution may therefore be expressed as

$$F(z) = A(1+z)^{-\nu/2} I_\nu(\zeta) + B(1+z)^{-\nu/2} K_\nu(\zeta), \tag{3.13}$$

where A and B are constants. However, the second part of this solution (the K_ν term) ought to be discarded. Consider the definition of $F(z)$ according to Equation (3.4). The gradient of $F(z)$ at $z = -1$ will be

$$\left.\frac{\mathrm{d}F}{\mathrm{d}z}\right|_{z=-1} = \sum_{N=0}^{\infty} N(-1)^{N-1} W(N) \leq \sum_{N=0}^{\infty} NW(N) = \bar{N} \tag{3.14}$$

and also

$$\left.\frac{\mathrm{d}F}{\mathrm{d}z}\right|_{z=-1} \geq -\sum_{N=0}^{\infty} NW(N) = -\bar{N}. \tag{3.15}$$

Hence the gradient is strictly finite at $z = -1$. Any explicit solution for $F(z)$ must exhibit this feature.

For convenience, the full solution expressed in Eq. (3.13) can be written as

$$F(z) = A(1+z)^{-\nu/2} I_\nu\left(C(1+z)^{1/2}\right) + B(1+z)^{-\nu/2} K_\nu\left(C(1+z)^{1/2}\right), \tag{3.16}$$

where A and B are integration constants and C is a specified combination of the parameters V, j_A and κ. The parameter ν takes values in the range $-1 \leq \nu \leq \infty$.

Let us first examine the I_ν part of the solution and write

$$F_I(z) = (1+z)^{-\nu/2} I_\nu\left(C(1+z)^{1/2}\right), \tag{3.17}$$

so using the identities (A.9) from the Appendix A for the derivative of Bessel function, we obtain

$$\begin{aligned} \frac{\mathrm{d}F_I}{\mathrm{d}z} &= -\frac{\nu}{2}(1+z)^{-\nu/2-1} I_\nu(\zeta) + (1+z)^{-\nu/2}\left(I_{\nu+1}(\zeta) + \frac{\nu}{\zeta} I_\nu(\zeta)\right)\frac{1}{2} C(1+z)^{-1/2} \\ &= \frac{C}{2}(1+z)^{-(\nu+1)/2} I_{\nu+1}(\zeta), \end{aligned} \tag{3.18}$$

where $\zeta = C(1+z)^{1/2}$.

Consider the behaviour of this gradient when $z \to -1$ and hence $\zeta \to 0$. Using the standard result in Equation (A.5), we find that as $z \to -1$,

$$\frac{\mathrm{d}F_I}{\mathrm{d}z} \sim \left(\frac{C}{2}\right)^{\nu+1} \frac{1}{\Gamma(\nu+2)}(1+z)^{-(\nu+1)/2}(1+z)^{(\nu+1)/2}, \tag{3.19}$$

which is finite, irrespective of the sign of ν.

Now let us examine the K_ν part of the solution in Eq. (3.16) and as before, we may write

$$F_K(z) = (1+z)^{-\nu/2} K_\nu \left(C(1+z)^{1/2}\right). \tag{3.20}$$

In order to evaluate the gradient of F_K we use the standard result (A.10), which then leads to

$$\frac{\mathrm{d}F_K}{\mathrm{d}z} = -\frac{C}{2}(1+z)^{-(\nu+1)/2} K_{\nu+1}(\zeta). \tag{3.21}$$

To see the behaviour of $\mathrm{d}F_K/\mathrm{d}z$ in the limit $z \to -1$ ($\zeta \to 0$), we apply the identity (A.7) together with (A.2) to Eq. (3.21), upon which we find that as $z \to -1$,

$$\frac{\mathrm{d}F_K}{\mathrm{d}z} = -\left(\frac{C}{2}\right)^{-\nu} \frac{\Gamma(\nu+1)}{2}(1+z)^{-(\nu+1)/2}(1+z)^{-(\nu+1)/2} \tag{3.22}$$

which goes to infinity even if ν is negative, remembering the fact that the value of ν has a lower limit of -1. This will violate the requirement that the gradient of $F(z=-1)$ must be finite. Hence the K_ν part of the solution (second term) in Equation (3.13) is untenable and we conclude that the actual solution to Equation (3.12) is

$$F(z) = A(1+z)^{-\nu/2} I_\nu(\zeta). \tag{3.23}$$

The normalisation constant A in this equation can be evaluated by considering Eq. (3.4) at $z = 1$ so that

$$F(1) = \sum_{N=0}^{\infty} W(N) = 1.$$

Hence

$$A = \frac{1}{2^{-\nu/2} I_\nu \left(2\sqrt{2X(\nu+1)}\right)} \tag{3.24}$$

and the solution to Eq. (3.5) is given by

$$F(z) = \left(\frac{1+z}{2}\right)^{-\nu/2} \frac{I_\nu \left(2\sqrt{X(\nu+1)(1+z)}\right)}{I_\nu \left(2\sqrt{2X(\nu+1)}\right)}. \tag{3.25}$$

Since all the individual parameters contributing to the definitions of X and ν are always positive, $X \in [0, \infty]$ and $\nu \in [-1, \infty]$. In the stochastic limit, where the gas-phase concentration of molecules and the particle size are both likely to be extremely small, one expects both X and ν to be at the lower end of the allowed ranges.

3.3.1.3 Mean population of reactants

In the stochastic context, the measurable mean adsorbed population, $\bar{N}$, is defined as

$$\bar{N} = \sum_{N=0}^{\infty} NW(N) = \left.\frac{\mathrm{d}F}{\mathrm{d}z}\right|_{z=1}. \tag{3.26}$$

Equation (3.23) can be differentiated to give

$$\begin{aligned}
\frac{\mathrm{d}F}{\mathrm{d}z} &= -A\frac{\nu}{2}(1+z)^{-\nu/2-1} I_\nu(\zeta) + A(1+z)^{-\nu/2} \frac{\mathrm{d}I_\nu}{\mathrm{d}\zeta}\frac{\mathrm{d}\zeta}{\mathrm{d}z} \\
&= -A\frac{\nu}{2}(1+z)^{-\nu/2-1} I_\nu(\zeta) + A(1+z)^{-\nu/2} \left(I_{\nu+1}(\zeta) + \frac{\nu}{\zeta} I_\nu(\zeta)\right) \frac{1}{2} C(1+z)^{-1/2} \\
&= A\frac{C}{2}(1+z)^{-(\nu+1)/2} I_{\nu+1}(\zeta),
\end{aligned} \tag{3.27}$$

where $C = 2\sqrt{X(\nu+1)}$. Evaluating this derivative at $z = 1$ and substituting the full expressions for A and C leads to the result

$$\bar{N} = \sqrt{\frac{X(\nu+1)}{2}} \frac{I_{\nu+1}\left(2\sqrt{2X(\nu+1)}\right)}{I_\nu \left(2\sqrt{2X(\nu+1)}\right)}. \tag{3.28}$$

3.3.1.4 Dimer production rate

The average production rate of the dimer C is defined as

$$j_C = \frac{\kappa}{2V} \sum_{N=0}^{\infty} N(N-1)W(N) = \frac{\kappa}{2V} \left. \frac{\mathrm{d}^2 F}{\mathrm{d}z^2} \right|_{z=1}. \tag{3.29}$$

This can be evaluated by differentiating Equation (3.27) and using identities (A.9) again. Two recurrence relationships for $I_{p-2}(r)$ and $I_{p+2}(r)$ are needed, which can be easily obtained from the original identities (A.9). We thus get the expression

$$\left. \frac{\mathrm{d}^2 F}{\mathrm{d}z^2} \right|_{z=1} = \frac{X(\nu+1)}{2} \frac{I_{\nu+2}\left(2\sqrt{2X(\nu+1)}\right)}{I_\nu\left(2\sqrt{2X(\nu+1)}\right)} \tag{3.30}$$

and hence we arrive at

$$j_C = \frac{j_A}{2} \frac{I_{\nu+2}\left(2\sqrt{2X(\nu+1)}\right)}{I_\nu\left(2\sqrt{2X(\nu+1)}\right)}. \tag{3.31}$$

Mass conservation in the steady state requires that the production rate of species C must be proportional to the difference between the molecules' incoming rate and the escape rate, namely

$$j_C = \frac{j_A - \lambda_A \bar{N}}{2}. \tag{3.32}$$

The factor 2 here reflects the fact that one molecule of species C is created at the cost of two A-molecules. It is in fact possible to show that the result (3.31) together with Equation (3.28) indeed satisfies this simple relationship.

3.3.1.5 Comparison with classical theory

The stochastic results obtained above must be compared with the corresponding classical results. In order to obtain the classical mean population of reactant molecules in a steady state, we may set $\mathrm{d}\bar{N}/\mathrm{d}t = 0$ in Eq. (3.3).

This leaves a simple quadratic equation in $\bar{N}$, the valid solution to which is

$$\begin{aligned}\bar{N}_{\text{classical}} &= \frac{-\lambda_A + \left(\lambda_A^2 + 4\frac{\kappa}{V} j_A\right)^{1/2}}{2\kappa/V} \\ &= \left[\left(\frac{\nu+1}{4}\right)^2 + \left(\frac{X(\nu+1)}{2}\right)\right]^{\frac{1}{2}} - \left(\frac{\nu+1}{4}\right). \end{aligned} \tag{3.33}$$

The classical expression for the dimer production rate is

$$j_C^{\text{classical}} = \frac{\kappa}{2V}\bar{N}^2_{\text{classical}}. \tag{3.34}$$

The stochastic results should agree with classical theory in three limits on the (X, ν) parameter space: when X is either very large or close to zero and when ν is very large. For example, when $X \to 0$, either the adsorption rate j_A is very small or the desorption rate λ_A is very large, both of which mean that the probability of finding an A-molecule on the surface at a given time is vanishingly small. This means the Poisson distribution will have the shape of an extremely narrow spike located almost at $N = 0$, which is an indication of the deterministic limit being re-established. When $X \to \infty$ for fixed ν, the adsorption rate overwhelms the desorption rate. The mean adsorbed population becomes very large, the probability distribution takes a Gaussian shape and the Poisson distribution is not required. This is clearly a regime where the stochastic approach should correspond with the classical treatment. Nevertheless, if the value of ν is simultaneously small, for example due to the reaction rate κ being large, it may complicate matters and the classical result may not necessarily be adequate enough. Finally, the limit $\nu \to \infty$ at fixed X corresponds to a situation where the adsorbed A-molecules evaporate so fast that there would be hardly any chance for two of them to react. Again, this means the Poisson distribution taking the shape of a quasi-delta function near $N = 0$ and both treatments should agree that the C-molecule production rate falls to zero.

$X \to \infty$, fixed $\nu \neq -1$

We check consistency in this limit between the two models using the standard approximation (A.6) in Equation (3.28) and then Taylor expanding the term arising in the denominator. Ignoring terms in $1/X$ and higher order enables one to show that

$$\bar{N} \approx \left(\frac{X(\nu+1)}{2}\right)^{\frac{1}{2}} - \frac{\nu}{4} - \frac{1}{8} \quad \text{as } X \to \infty. \tag{3.35}$$

To check this against the classical result, the square root term in Eq. (3.33) can again be Taylor expanded about the point $X \to \infty$ and then if we ignore the terms in $1/X$, we obtain

$$\bar{N}_{\text{classical}} \approx \left(\frac{X(\nu+1)}{2}\right)^{\frac{1}{2}} - \frac{\nu}{4} - \frac{1}{4} \quad \text{as } X \to \infty. \tag{3.36}$$

Equations (3.35) and (3.36) indicate that the stochastic model yields approximately the same result as the classical kinetic theory does in the asymptotic limit of very large flux of reacting molecules. Note, however, that if $\nu < 0$, and particularly if it is close to -1, it may act to neutralise the effect of large X in Eqs. (3.35) and (3.36). In that case, some noticeable difference may appear between the stochastic and classical results.

$\nu \to \infty$, fixed X

A second limit where the stochastic model converges with its classical counterpart is when $\nu \to \infty$ at constant X. Although ν appears as part of the argument of Bessel function in Equation (3.28), it is unwise to employ the approximation (A.6) since ν is also the order of the Bessel function. A better approach would be to consider the series expansion of $\bar{N}$ in Eq. (3.28) through Equation (A.3) so that

$$\bar{N} = \sqrt{\frac{X(\nu+1)}{2}}\sqrt{2X(\nu+1)}\,\frac{\Gamma(\nu+1)}{\Gamma(\nu+2)}\,\frac{\left[1 + \frac{2X(\nu+1)}{(\nu+2)} + \frac{1}{2}\frac{4X^2(\nu+1)^2}{(\nu+3)(\nu+2)} + \cdots\right]}{\left[1 + \frac{2X(\nu+1)}{(\nu+1)} + \frac{1}{2}\frac{4X^2(\nu+1)^2}{(\nu+2)(\nu+1)} + \cdots\right]} \tag{3.37}$$

For a relatively large value of ν, assuming that $\nu + 1 \approx \nu + 2$ and so on, the two squared braces above would cancel each other out. Noting that $\Gamma(\nu + 2) = (\nu + 1)\,\Gamma(\nu + 1)$, the approximation for $\bar{N}$ is

$$\bar{N} \approx X. \tag{3.38}$$

On the other hand, if we write the classical expression in Equation (3.33) as

$$\bar{N}_{\text{classical}} = \left(\frac{\nu+1}{4}\right)\left[1 + \frac{8X}{\nu+1}\right]^{\frac{1}{2}} - \left(\frac{\nu+1}{4}\right) \tag{3.39}$$

then the series expansion for $\bar{N}_{\text{classical}}$ may be written as

$$\begin{aligned} \bar{N}_{\text{classical}} &= \left(\frac{\nu+1}{4}\right)\left[1 + \frac{4X}{\nu+1} + \cdots\right] - \left(\frac{\nu+1}{4}\right) \\ &\approx X + O\left(\frac{1}{\nu+1}\right) \end{aligned} \tag{3.40}$$

for large ν. Hence the classical and stochastic models yield approximately the same production rate when $\nu \to \infty$.

<u>$X \to 0,\ \nu \neq -1$</u>

For this limit we use the approximation (A.5) in Equation (3.28) together with the identity (A.4). It can then be easily shown that

$$\begin{aligned} \bar{N} &\approx \sqrt{\frac{X(\nu+1)}{2}}\sqrt{2X(\nu+1)}\,\frac{\Gamma(\nu+1)}{(\nu+1)\Gamma(\nu+1)} \\ &\approx X \\ &= 0 \quad \text{as } X \to 0. \end{aligned} \tag{3.41}$$

To compare this with the classical result, let us write Equation (3.33) as

$$\bar{N}_{\text{classical}} = (\nu+1)^{1/2}\left[\frac{\nu+1}{16} + \frac{X}{2}\right]^{\frac{1}{2}} - \left(\frac{\nu+1}{4}\right), \tag{3.42}$$

so that in the squared brackets the terms in X and ν are separated. When X is very small, this can be Taylor expanded to obtain the series

$$\begin{aligned}\bar{N}_{\text{classical}} &= (\nu+1)^{1/2}\left[\left(\frac{\nu+1}{16}\right)^{\frac{1}{2}}+\frac{X}{4}\left(\frac{\nu+1}{16}\right)^{-\frac{1}{2}}-\frac{X^2}{32}\left(\frac{\nu+1}{16}\right)^{-\frac{3}{2}}+\cdots\right] \\ &\qquad -\left(\frac{\nu+1}{4}\right) \\ &\approx \left[\frac{\nu+1}{4}+X-\frac{2X^2}{\nu+1}\right]-\left(\frac{\nu+1}{4}\right) \\ &= X-\frac{2X^2}{\nu+1}\,. \end{aligned} \tag{3.43}$$

Hence the classical result also yields a zero reactant population when $X = 0$, but as $X \to 0$, $\bar{N}_{\text{classical}}$ approaches the zero in a different manner than the stochastic expression (3.41). This will be evident when we look at some of the results in §3.5.1.1.

$\nu \to -1$, fixed X

This is not a limit where stochastic and classical treatments coincide, but it is included for an interesting comparison with the model of Mozurkewich [38]. The limit $\nu \to -1$ at fixed X corresponds to κ/V being very much greater than λ_A. In this case the series expansion appearing in Equation (3.37) may be written as

$$\bar{N} = X(\nu+1)\left(\frac{\Gamma(\nu+1)}{(\nu+1)\Gamma(\nu+1)}\right)\frac{1}{(1+2X)}\frac{\left[1+\frac{2X\epsilon}{\nu+2}+O(\epsilon^2)\right]}{\left[1+\frac{2X^2\epsilon}{\nu+2}+O(\epsilon^2)\right]} \tag{3.44}$$

where $\epsilon = \nu + 1$ and is assumed to be very small as $\nu \to -1$. If we ignore the terms of the order of ϵ^2 and higher and perform a Taylor expansion of the squared brace in the denominator, we arrive at

$$\bar{N} = \frac{X}{1+2X}\left[1+\frac{2X(1-X)}{(\nu+2)}\epsilon+O(\epsilon^2)\right]. \tag{3.45}$$

If ϵ is very small ($\nu \to -1$), the factor in squared brace can be neglected. If we now consider large X, then $\bar{N} \sim 0.5$, which is in fine agreement with the result obtained by Mozurkewich [38] for atmospheric species in the case of $\kappa/V \to \infty$ with $\lambda_A = 0$. However if X is very small, Equation (3.45) suggests that $\bar{N} \sim 0$, which is consistent

with what one expects from the classical result (3.33).

3.3.2 Reaction A + B → C

We now extend the problem to the situation where molecules of two different chemical species A and B are incident onto the particle surface and occasionally react with each other to form some product C. In such a case, we need to consider the probability $W(N_A, N_B; t)$ of finding exactly N_A molecules of species A and N_B molecules of species B on the particle. Similarly, we shall now have two different rates of adsorption of molecules, j_A and j_B, and for molecular escape, we introduce rate coefficients λ_A and λ_B, corresponding to the two species A and B respectively. The master equation for such a process will be much more complicated in this case:

$$\begin{aligned}\frac{\mathrm{d}W(N_A,N_B;t)}{\mathrm{d}t} &= j_A[W(N_A-1,N_B;t)-W(N_A,N_B;t)]\\ &\quad + j_B[W(N_A,N_B-1;t)-W(N_A,N_B;t)]\\ &\quad + \frac{\kappa}{V}[(N_A+1)(N_B+1)W(N_A+1,N_B+1;t)\\ &\qquad - N_AN_BW(N_A,N_B;t)]\\ &\quad + \lambda_A[(N_A+1)W(N_A+1,N_B;t)-N_AW(N_A,N_B;t)]\\ &\quad + \lambda_B[(N_B+1)W(N_A,N_B+1;t)-N_BW(N_A,N_B;t)]. \quad (3.46)\end{aligned}$$

This equation is constructed using the same kind of arguments as before. Once again, in order to simplify this equation we introduce a generating function

$$F(x,y;t)=\sum_{\substack{N_A=0\\N_B=0}}^{\infty} W(N_A,N_B;t)x^{N_A}y^{N_B} \quad (3.47)$$

so that the master equation (3.46) reduces to a second order time-dependent partial differential equation:

$$\frac{\partial F}{\partial t} = \left(j_A(x-1) + j_B(y-1) - \lambda_A(x-1)\frac{\partial}{\partial x} - \lambda_B(y-1)\frac{\partial}{\partial y} - \frac{\kappa}{V}(xy-1)\frac{\partial^2}{\partial x \partial y} \right) F. \quad (3.48)$$

Similarly to the previous case, the average numbers of molecules A and B are defined as

$$\begin{aligned} \bar{N}_A &= \sum_{\substack{N_A=0 \\ N_B=0}}^{\infty} N_A W(N_A, N_B; t) = \left.\frac{\partial F}{\partial x}\right|_{x=y=1} \\ \bar{N}_B &= \sum_{\substack{N_A=0 \\ N_B=0}}^{\infty} N_B W(N_A, N_B; t) = \left.\frac{\partial F}{\partial y}\right|_{x=y=1}. \end{aligned} \quad (3.49)$$

This will replace the classical rate equations which, under equilibrium, would have been the coupled equations

$$\begin{aligned} 0 &= j_A - \frac{\kappa}{V}\bar{N}_A\bar{N}_B - \lambda_A\bar{N}_A \\ 0 &= j_B - \frac{\kappa}{V}\bar{N}_A\bar{N}_B - \lambda_B\bar{N}_B. \end{aligned} \quad (3.50)$$

Unlike the previous case of A+A $\rightarrow$ C, Eq. (3.48) cannot be solved analytically. However, it is possible to gain some insight about the approximate nature of its solution. If we ignore the last term in Eq. (3.48), then by guessing that

$$F = \exp[\bar{N}_A(x-1) + \bar{N}_B(y-1)] \quad (3.51)$$

it is possible to get back Eqs. (3.50) from Eq. (3.48). Indeed, this is the exact solution in the absence of any reactions taking place on the grain surface (i.e. $\kappa = 0$), with $\bar{N}_A = j_A/\lambda_A$ and $\bar{N}_B = j_B/\lambda_B$. It represents independent Poisson distributions for the two populations, whose covariance is zero, so that the deterministic rate law should hold for all reactions. We shall use this generating function, with $\bar{N}_A$ and $\bar{N}_B$ as given above, to provide a starting point for the computations described in §3.4.2.

3.4 Numerical computation

3.4.1 $A + A \rightarrow C$

The stochastic result obtained in the above model for the simple reaction $A+A \rightarrow C$ was compared with that of a Monte Carlo simulation, in which we provide j_A, κ, λ_A and V as input parameters to begin with. Given these conditions, the time evolution of the system is then observed until it settles into a steady state.

After starting the time-loop, at each time-step we define three relative probabilities

$$\begin{aligned} p_1 &= j_A \\ p_2 &= \lambda_A N \\ p_3 &= \frac{\kappa}{2V} N(N-1), \end{aligned} \tag{3.52}$$

which respectively represent the reactant molecules' influx on the particle surface, the escape of a molecule from the surface and the reaction between two molecules in order to produce a dimer. These relative probabilities need to be normalised so that they lie in the range [0, 1]. To do this, we weight them with a specific number μ, i.e. evaluate

$$p_i' = \frac{p_i}{\mu} \quad (i = 1, 2, 3). \tag{3.53}$$

In practice, the number μ is chosen by trial and error such that $p_1' + p_2' + p_3'$ never overshoots unity. μ is input at the very beginning of the simulation and remains unaltered during the rest of the procedure. Often a sufficiently safe guess for μ is to take it equal to the value of $p_1 + p_2 + p_3$ based on the anticipated maximum value of N. Nevertheless this scheme will almost invariably leave $p_1' + p_2' + p_3'$ slightly short of unity. This shortfall is assigned as p_4', the probability corresponding to 'no event'

taking place on the grain surface (see Figure 3.3). The normalised probabilities p'_1, p'_2 and p'_3 now represent the probabilities of adsorption, escape or reaction of molecules respectively, such that $(p'_1 + p'_2 + p'_3 + p'_4) = 1$. The introduction of p'_4 is essentially due to the fact that N varies as the simulation progresses. In practice it can be kept very small compared to p'_1, p'_2 and p'_3 if μ is chosen wisely enough, so as to speed up the simulation. Though it is also possible to eliminate p'_4 entirely [33, 34].

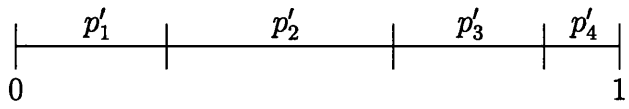

Figure 3.3: The normalised probabilities corresponding to the three possible events (p'_1, p'_2 and p'_3) plus the probability of no event taking place (p'_4). These are such that they add up to 1 in order to allow a Monte Carlo simulation.

Since each one of p'_1, p'_2 and p'_3 will represent a unique segment in the interval [0, 1], the usual Monte Carlo procedure for simulating the three possible events at each time-step can be employed. We continue for a time-span that is long enough to produce a steady state mean population. The results are shown in §3.5.1.1.

3.4.2 $A + B \rightarrow C$

The solution of Eq. (3.48) promises to provide all the information about the system of our interest. In the absence of an analytic solution, some kind of numerical technique should be sought. In the following subsections, we explore three possible ways of solving Eq. (3.48) numerically. The first approach is a naïve one that occurs

most naturally to one's mind, but is not very reliable for large mean population of reactants. Approach 2 is a modification of the first approach, taking into account the shortcomings of the Approach 1. For not-too-small values of the mean population, the second approach is valuable, but tends to be very unstable for very small $\bar{N}_A$ or $\bar{N}_B$. Hence a third technique, which is ideal for low reactant populations, is described at the end of this Section.

3.4.2.1 Approach 1

The simplest method to solve Eq. (3.48) is to discretise the (x, y) space and apply the Equation at each point on the discretised space in the steady state conditions, i.e. with $\partial F/\partial t = 0$. Let us divide the (x, y) space in such a way that a specified interval on x axis (e.g. $x = 0 \ldots 1$) is divided into n divisions of length Δx and a given interval on y axis is divided into n divisions of length Δy (Figure 3.4). The co-ordinates of each point on the grid may be assigned as

$$\begin{aligned} x_i &= i\Delta x \quad (i = 0, 1, 2, \ldots n) \\ y_j &= j\Delta y \quad (j = 0, 1, 2, \ldots n). \end{aligned} \tag{3.54}$$

Let us further simplify the notation by calling

$$F(x_i, y_j) \equiv F_{i,j}$$

and so on. Provided that Δx and Δy are very small, Eq. (3.48) for the points *not* on the boundary of the defined grid may be approximated as

$$\begin{aligned} 0 \;=\; & j_A(x_i - 1)F_{i,j} + j_B(y_j - 1)F_{i,j} \\ & - \lambda_A(x_i - 1)\left[\frac{F_{i+1,j} - F_{i-1,j}}{2\Delta x}\right] - \lambda_B(y_j - 1)\left[\frac{F_{i,j+1} - F_{i,j-1}}{2\Delta y}\right] \\ & - \frac{\kappa}{V}(x_i y_j - 1)\left[\frac{F_{i+1,j+1} - F_{i+1,j-1} - F_{i-1,j+1} + F_{i-1,j-1}}{(2\Delta x)(2\Delta y)}\right] \end{aligned} \tag{3.55}$$

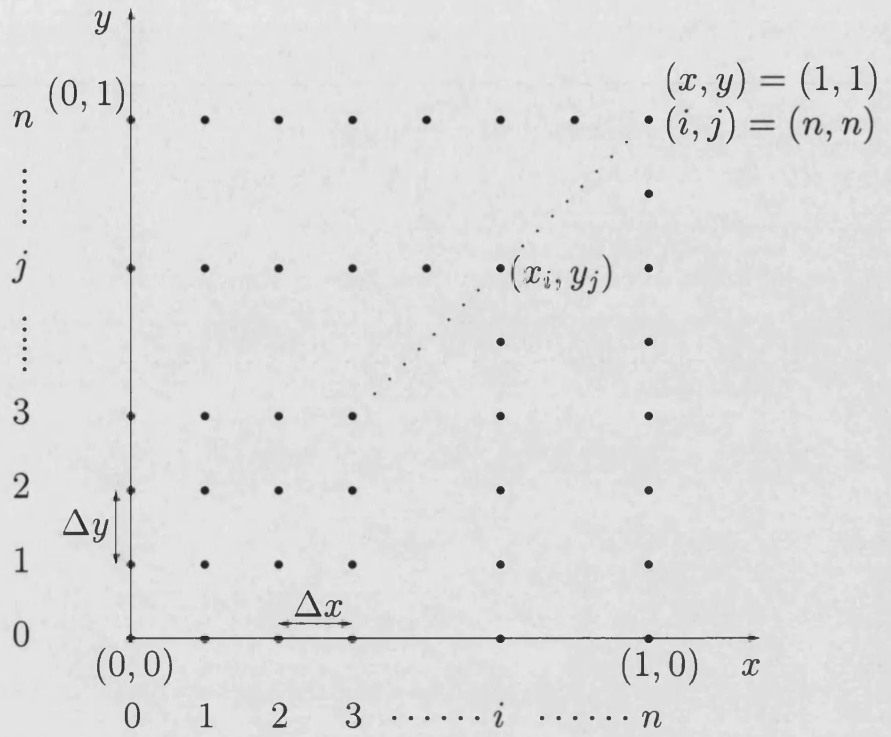

Figure 3.4: Discretization of the (x, y) space appearing in the generating function $F(x, y)$.

using approximate expressions for the derivatives of F with respect to x and y. For the 'boundary points' on the grid, slightly different approximations for the derivatives need to be employed by considering only the points within the defined grid.

We need to determine the numerical value of F at each point on the discretised (x, y) space of interest. Thus a grid of $(n + 1) \times (n + 1)$ points will pose $(n + 1)^2$ unknowns to be evaluated in general. With Eq. (3.55) applied to each of these $(n + 1)^2$ points on the grid, we have as many equations in hand, leading to a system of linear equations of the form

$$
\begin{aligned}
s_{1,1}F_{0,0} + s_{1,2}F_{0,1} + \cdots + s_{1,n^2+2n}F_{n,n-1} + s_{1,(n+1)^2}F_{n,n} &= 0 \\
s_{2,1}F_{0,0} + s_{2,2}F_{0,1} + \cdots + s_{2,n^2+2n}F_{n,n-1} + s_{2,(n+1)^2}F_{n,n} &= 0 \\
\cdots\cdots\cdots

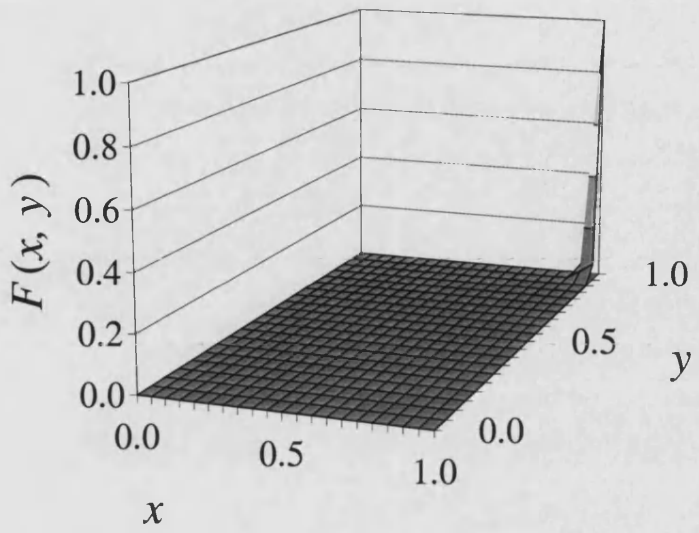

Figure 3.5: The function $F(x, y)$ for a typical set of parameters.

Figure 3.5 shows one such solution achieved with this method with a set of parameters that are characteristic of conditions found in interstellar space or the Earth's upper atmosphere where gas concentrations are very low. The striking feature of this plot is that $F(x, y)$ rises extremely sharply near the point of our interest, that is at $x = y = 1$, where the derivatives of $F(x, y)$ are to be calculated to obtain $\bar{N}_A$ and $\bar{N}_B$ according to Equation (3.49). Given the discretised nature of the (x, y) space, this is problematic. No matter how finely we discretise the (x, y) space, the accuracy of the gradient at $x = y = 1$ cannot be guaranteed. For small gradients, however, this technique will be quite powerful, requiring negligible computational time.

3.4.2.2 Approach 2

Due to such problems encountered in dealing with $F(x,y)$, a safer strategy is to work with the function

$$G(x,y;t) = \ln[F(x,y;t)] \tag{3.58}$$

which is likely to be better behaved than $F(x,y)$ at $x = y = 1$. Equation (3.48) can be re-written in terms of $G(x,y;t)$ as

$$\begin{aligned}\frac{\partial G}{\partial t} &= j_A(x-1) + j_B(y-1) - \lambda_A(x-1)\frac{\partial G}{\partial x} - \lambda_B(y-1)\frac{\partial G}{\partial y} \\ &\quad - \frac{\kappa}{V}(xy-1)\left\{\frac{\partial^2 G}{\partial x \partial y} + \left(\frac{\partial G}{\partial x}\right)\left(\frac{\partial G}{\partial y}\right)\right\}.\end{aligned} \tag{3.59}$$

Furthermore, using the identity $\partial \ln F/\partial x = (1/F)(\partial F/\partial x)$, and using the fact that $F(1,1;t) = 1$, we note that

$$\bar{N}_A = \left.\frac{\partial F}{\partial x}\right|_{x=y=1} = F\left.\frac{\partial \ln F}{\partial x}\right|_{x=y=1} = \left.\frac{\partial G}{\partial x}\right|_{x=y=1} \tag{3.60}$$

and similarly $\bar{N}_B = \partial G/\partial y|_{x=y=1}$.

Here matrix inversion is no longer an option as it was in the first method, since the discretised equations would be non-linear. In order to solve Equation (3.59), j_A, j_B, κ, λ_A, λ_B and V are provided as input parameters and the time dependent problem is solved, starting from a trial solution and terminating when the solution has converged to a steady state. As a first approximation, we make use of Eq. (3.51) and thus assign the following values to G and to its first and second derivatives at time $t = 0$:

$$G(x,y;t=0) = \bar{N}_A^{\text{clas}}(x-1) + \bar{N}_B^{\text{clas}}(y-1) \tag{3.61}$$

$$\left.\frac{\partial G}{\partial x}\right|_{t=0} = \bar{N}_A^{\text{clas}}, \qquad \left.\frac{\partial G}{\partial y}\right|_{t=0} = \bar{N}_B^{\text{clas}} \tag{3.62}$$

$$\left.\frac{\partial^2 G}{\partial x \partial y}\right|_{t=0} = 0 \tag{3.63}$$

where $\bar{N}_A^{\text{clas}} = j_A/\lambda_A$ and $\bar{N}_B^{\text{clas}} = j_B/\lambda_B$, which are the classical values of $\bar{N}_A$ and $\bar{N}_B$ respectively in the absence of any reactions ($\kappa = 0$) as determined by Eqs. (3.50). At $t = 1$, the reaction rate constant κ is then 'switched on', which enables one to allow $G(x, y; t)$ to evolve in time according to Eq. (3.59) while maintaining the boundary condition $G(x = y = 1; t) = 0$.

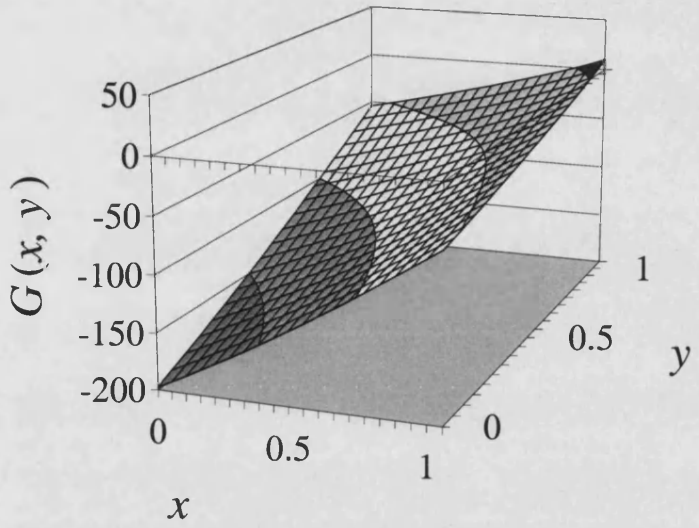

Figure 3.6: The function $G(x, y)$ for a typical set of parameters when equilibrium has been reached at the end of time-evolution.

For computational purpose, the (x, y) space as well as the time t are discretised and thereafter the Explicit Euler Method is applied in order to obtain an iterative solution over the time-series. If the time-steps are labelled by m and if we call $G(x, y; t_m) = G_m$, then the left hand side of Eq. (3.59) may be expressed as

$$\frac{\partial G(x, y; t_m)}{\partial t} = \frac{G_m - G_{m-1}}{\Delta t} \tag{3.64}$$

where Δt should be a very small time-step size, and is provided as an input. The right hand side of Eq. (3.59) is evaluated by considering the values of all the derivatives at the 'previous' time-step t_{m-1}, which are known from the previous iteration.

Thus Eq. (3.59) can be reduced to the approximation

$$G_m = G_{m-1} + \left[j_A(x-1) + j_B(y-1) - \lambda_A(x-1)\left.\frac{\partial G}{\partial x}\right|_{m-1} - \lambda_B(y-1)\left.\frac{\partial G}{\partial y}\right|_{m-1} - \frac{\kappa}{V}(xy-1)\left\{ \left.\frac{\partial^2 G}{\partial x \partial y}\right|_{m-1} + \left.\frac{\partial G}{\partial x}\right|_{m-1}\left.\frac{\partial G}{\partial y}\right|_{m-1} \right\} \right] \Delta t \qquad (3.65)$$

where the derivatives of G may be approximated using the discretised nature of the (x, y) space in the same was as in Eq. (3.55). This leaves only one unknown G_m, which can be explicitly evaluated. We thus proceed in the time-series until G settles into equilibrium, i.e. $G_m \simeq G_{m-1}$ for each point (x_i, y_j), eliminating the need for any further iterations. A typical converged solution is shown in Figure 3.6.

The evolution of G is a very delicate process and can diverge very quickly away from the actual equilibrium state, particularly as the mean populations become smaller. This arises from the insufficient means by which the solution is fully determined with the help of boundary conditions. To avoid this, it is best to keep Δt as small as possible. The explicit Euler method has been used here for simplicity, though it is possible to use an implicit technique such as the Crank-Nicholson method [46], which is far more stable but much more complex. Alternatively, for very small mean populations, a further approach can be examined.

3.4.2.3 Approach 3

Yet another method of solving the problem of the reaction A+B $\rightarrow$ C is to evolve the probabilities $W(N_A, N_B; t)$ numerically in time until a steady state has been reached. As with Approach 2, we may discretise time t and replace the dt in Eq. (3.46) with a very small Δt. After providing j_A, j_B, κ, λ_A, λ_B and V as input parameters we let the system evolve according to equation (3.46). The initial condition was chosen as $W(0, 0; t = 0) = 1$ and all other elements of W equal to zero, indicating that there

are no molecules on the particle surface to start with.

Ideally, one needs to evolve an infinite number of probability elements $W(N_A, N_B; t)$, where $N_A \in [0, \infty]$ and $N_B \in [0, \infty]$. However, this is not practicable in reality given the finite capacity of computers. As a way out of this difficulty, we may set an upper limit on the maximum number of A and B molecules that the particle surface can possess at a given time, namely $N_A^{\max}$ and $N_B^{\max}$, instead of infinity. Effectively, this means that we are approximating the mean populations in Equations (3.49) by

$$\begin{aligned} \bar{N}_A &= \sum_{N_A=0}^{N_A^{\max}} \sum_{N_B=0}^{N_B^{\max}} N_A W(N_A, N_B; t) \\ \bar{N}_B &= \sum_{N_A=0}^{N_A^{\max}} \sum_{N_B=0}^{N_B^{\max}} N_B W(N_A, N_B; t). \end{aligned} \tag{3.66}$$

Since we expect the probability distribution to look something like those in Figure 3.1, this clearly means that the values of $N_A^{\max}$ and $N_B^{\max}$ must be sufficiently large in order to allow the elements of W towards the end of the series in Equation (3.66) to be negligible. That will ensure that the achieved solution is satisfactory.

This technique works very well when the steady state corresponds to reasonably small $\bar{N}_A$ and $\bar{N}_B$ so that some modest values of $N_A^{\max}$ and $N_B^{\max}$ will suffice to obtain a good solution. However, as the mean populations reach high values, through increasing j_A or j_B for example, one needs to have quite large $N_A^{\max}$ and $N_B^{\max}$. Hence, depending on the available computer resources, there is an upper limit on the values of $\bar{N}_A$ and $\bar{N}_B$ that one can obtain with this technique. For large mean populations, Approach 2 would be more suitable. It is worth mentioning here that an alternative method for solving this problem was explored by Green *et al.* [44], in which the master equation (3.46) was solved directly in the steady state through a matrix inversion approach similar to that in Approach 1.

3.5 Results and discussion

An extreme example of grain surface chemical reactions taking place under ultra low concentration of both the gas and the dust is seen in the interstellar medium. For this reason the stochastic model developed in §3.3 was tested under parameters corresponding to interstellar conditions and by making use of the computational techniques described in §3.4. It was then applied to conditions more akin to the terrestrial atmosphere.

3.5.1 $A + A \rightarrow C$

3.5.1.1 Interstellar conditions

The simplest example of this type of reaction is the production of molecular hydrogen from its atomic state, i.e. $H + H \rightarrow H_2$. Although simple, this is a very significant reaction in the Universe. The result of the analytical stochastic approach, the Monte Carlo simulation, as well as the classical theory for this process under interstellar conditions is shown in Figures 3.7 and 3.8. The calculations have been performed for dust particles of radius ~ 0.1 μm within a cloud of gas phase concentration of 1 atom cm^{-3}, held at a temperature of 10 K. The corresponding rate parameters used here are $\kappa/(2V) = 5.1 \times 10^4$ s^{-1} and $\lambda_A = 1.9 \times 10^{-3}$ s^{-1}, taken from prior calculations of Caselli *et al.* [40]. A value of $j_A = 10^7 \text{s}^{-1}$ corresponds to $X = 5 \times 10^9$ and $j_A = 10^{-5}$ s^{-1} is equivalent to $X = 5 \times 10^{-3}$.

In Figure 3.7 the average population of A-molecules in equilibrium, $\bar{N}$, is plotted against the influx j_A. Figure 3.8 shows the production rate of C for the same set of parameters. Note that the stochastic $\bar{N}$ is higher than the classical counterpart, whereas the situation is reversed for j_C since according to Eq. (3.32), a higher $\bar{N}$

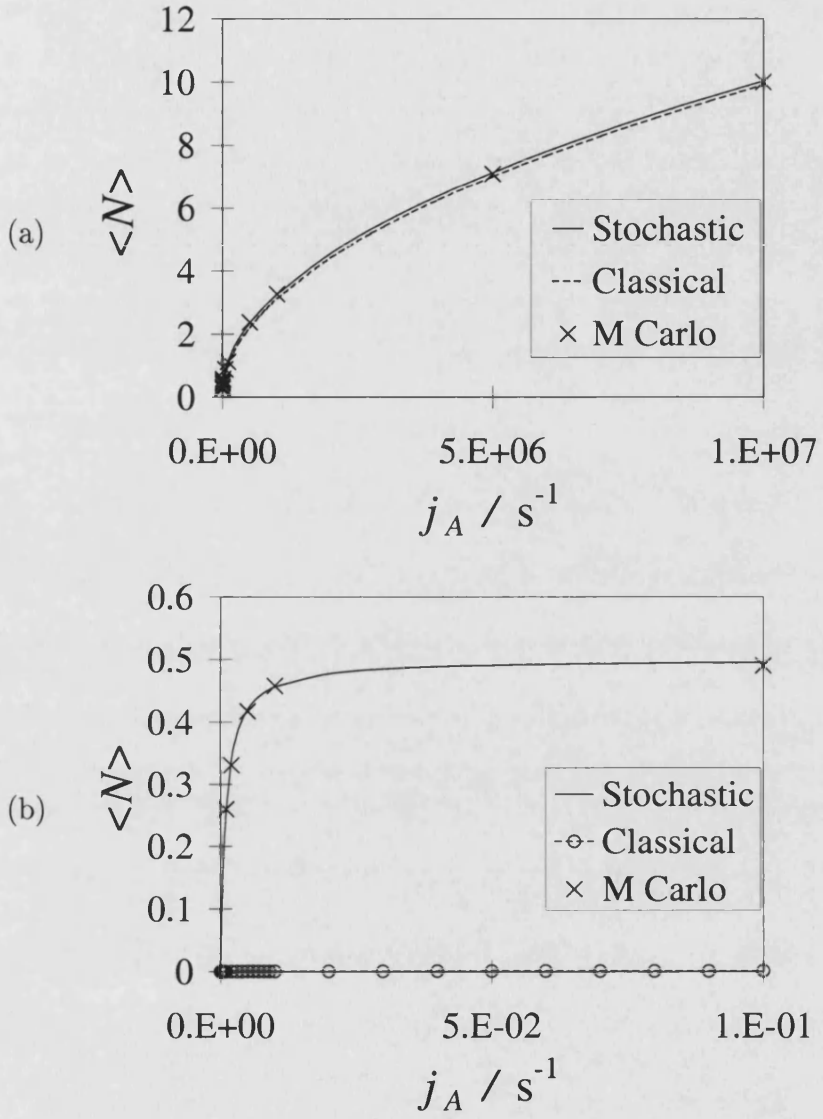

Figure 3.7: Average number of A-molecules (hydrogen atoms) present on the particle surface in equilibrium ($\bar{N}$, denoted here as $< N >$) as a function of flux j_A in interstellar conditions with 0.1 μm dust ($\lambda_A = 1.9 \times 10^{-3}$ s^{-1}). The corresponding $\nu = -0.99999...$ which is very close to the lower limit of -1. (a) The stochastic curve agrees with the classical curve at large j_A. (b) When we zoom in around ultra-small value of j_A, a considerable difference appears between the two. The stochastic curve is in good agreement with Monte Carlo calculations. Note the different scales on the abscissa.

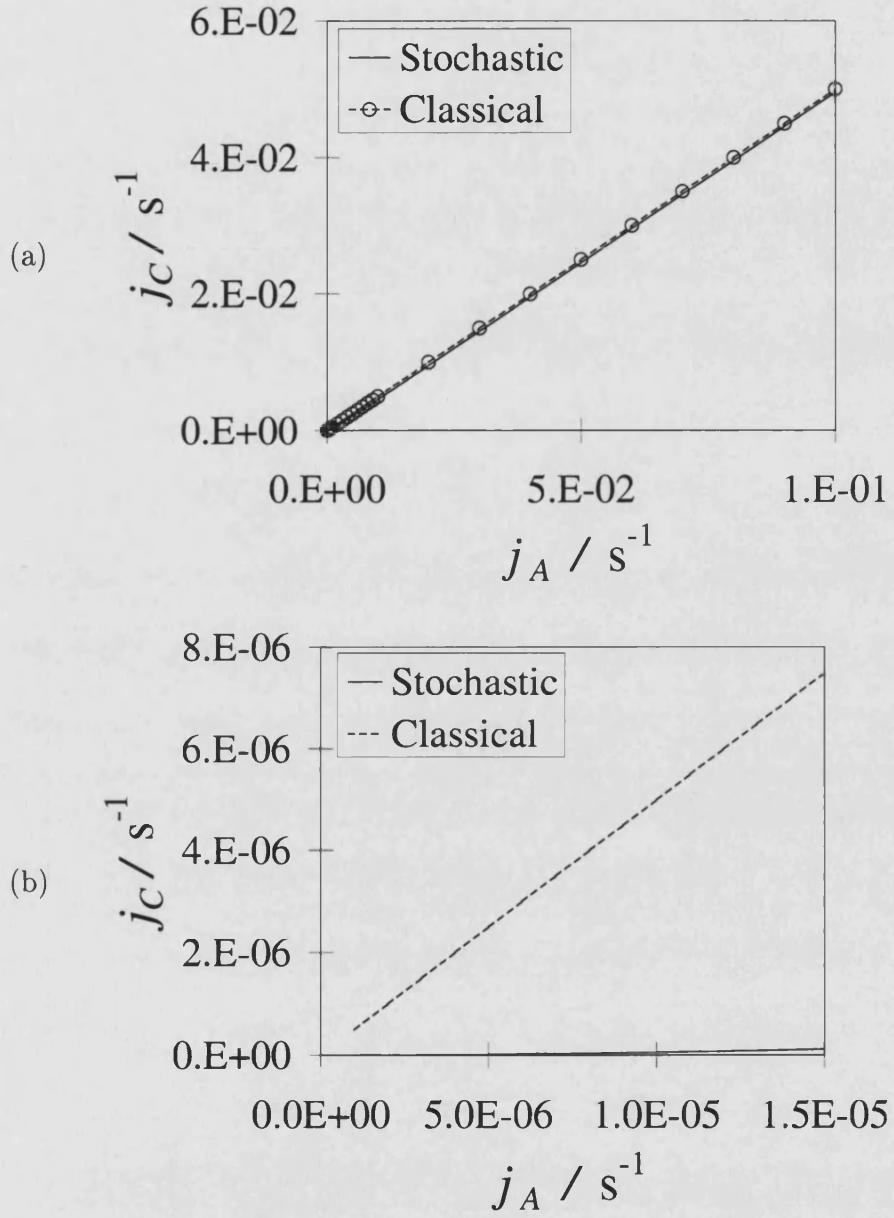

Figure 3.8: Production rate of H_2 (j_C) as a function of j_A in interstellar conditions with 0.1 μm radius dust and $\lambda_A = 1.9 \times 10^{-3}$ s^{-1} (corresponding to $\nu = -0.99999...$). (a) Stochastic model converges towards its classical counterpart at large flux. (b) When the flux is very small, the stochastic approach predicts such a small production rate that it can hardly be seen at the bottom of the graph.

means a lower j_C.

Clearly, Monte Carlo calculations are in good agreement with the stochastic model, both being significantly different from the classical curve at low flux or at low gas concentration. The value $j_A = 1.4 \times 10^{-5}\ \mathrm{s}^{-1}$ is what we typically expect in such interstellar clouds [40], for which we get $\bar{N}_{\mathrm{classical}} = 8.28 \times 10^{-6}$, in contrast to $\bar{N}_{\mathrm{stochastic}} = 7.27 \times 10^{-3}$. In terms of the dimer production rate, we have $j_C^{\mathrm{classical}} = 7 \times 10^{-6}\,\mathrm{s}^{-1}$ and $j_C^{\mathrm{stochastic}} = 1 \times 10^{-7}\,\mathrm{s}^{-1}$. Hence the classical rate equation approach may overestimate the production rate by over one order of magnitude.

Another feature that can be clearly seen in Figs. 3.7 and 3.8 is that the stochastic model converges with the classical theory at both limits of j_A, i.e. as $j_A \to 0$ and $j_A \to \infty$. There is a certain range of the parameter space for which the classical kinetics becomes unreliable; the range being characterised by the values of X and ν. The limits where classical results hold reasonable validity are discussed in §3.3.1.5.

Figures 3.9 and 3.10 show the area on the (X, ν) plane where the stochastic $\bar{N}$ and j_C differ from the classical results of rate equations. Figure 3.11 shows a part of the same data presented in Figure 3.10, but in a simpler way; here the ratio of production rates as predicted by the classical and stochastic models, namely $j_C^{\mathrm{clas}}/j_C^{\mathrm{stoch}}$, is plotted as a function of the parameter ν for three arbitrarily chosen values of X. Figures 3.10 and 3.11 show that the distinction between the dimer production rates according to the two models fades away with increasing X. Both the ratios $\bar{N}_{\mathrm{clas}}/\bar{N}_{\mathrm{stoch}}$ and $j_C^{\mathrm{clas}}/j_C^{\mathrm{stoch}}$ are 1 at large ν. In order to have a clear difference between the results of the two approaches, small values of both X and ν are required.

In the classical treatment, the rate of dimer formation is taken to be proportional

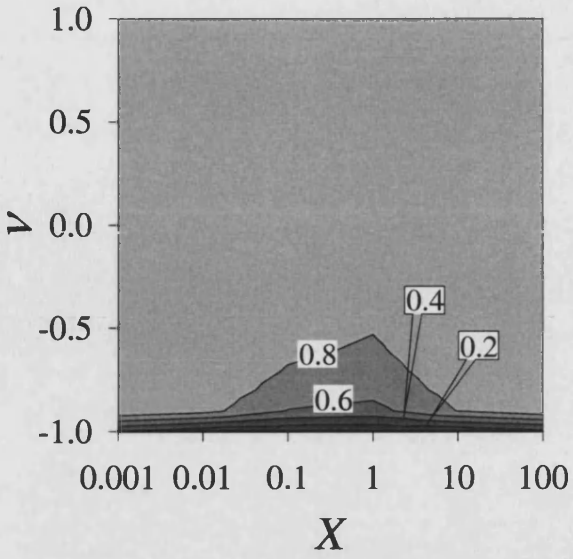

Figure 3.9: The contour plot of the ratio $\bar{N}_{\text{classical}}/\bar{N}_{\text{stochastic}}$ as a function of X and ν. The values quoted in white boxes refer to the boundaries between the shades. The ratio is zero at $\nu = -1$ for all X due to $\bar{N}_{\text{classical}}$ falling to zero. The stochastic model yields similar results to the classical rate equations in the three limits discussed in §3.3.1.5, so the ratio will reach 1 as $X \to 0$, $X \to \infty$ or $\nu \to \infty$.

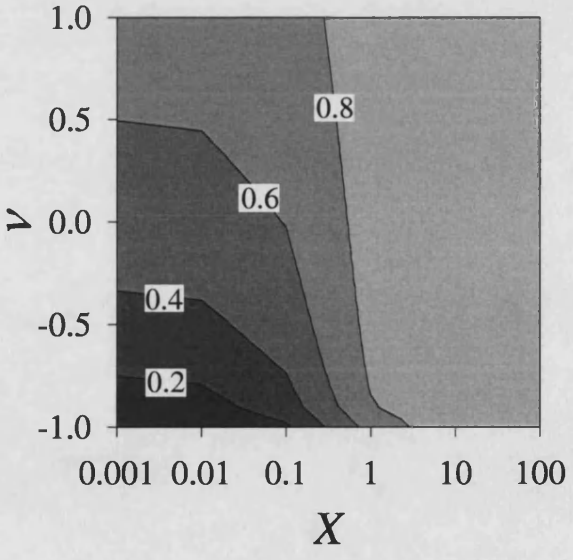

Figure 3.10: The contour plot of the ratio $j_C^{\text{stochastic}}/j_C^{\text{classical}}$ as a function of X and ν. This ratio is also zero at $\nu = -1$ where $j_C^{\text{classical}} \to \infty$ due to infinitely large κ.

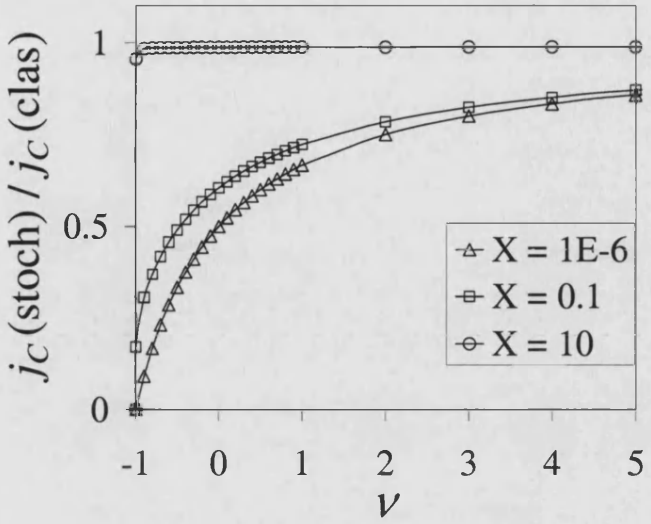

Figure 3.11: The ratio of dimer production rate as predicted by the stochastic model presented here, divided by the classical rate equations prediction, as a function of the parameter ν. The ratio of roughly 1 indicates that the two methods agree. The difference between the two models is visible when X and ν are small.

to $\bar{N}^2$. When the mean population of reactants is below unity, say $\bar{N} = 0.5$, the classical approach will calculate the dimer production rate as $\kappa/2V$ times 0.25. In reality, if the probability distribution of the reactant population exhibits a monotonic decrease such as in Figure 3.1 ($N \sim 1$ curve), then a mean population of 0.5 means that on most instances there will be either one monomer present on the surface, or none at all. None of these two situations will actually lead to dimer production, as at least two monomers are required to form a dimer. Only through a lucky chance there will be two or more monomers present on the surface simultaneously, which may then lead to dimer production. That would be indeed rare. The classical rate equation approach, which considers 'half' a molecule to be present on the surface at all times, overlooks the unbreakable nature of the reactant molecules and hence leads to an error in the dimer production rate when $\overline{N(N-1)} \neq \bar{N}^2$. In the first three limits discussed in §3.3.1.5, this error becomes negligible compared to the actual

mean population; hence the two models appear to agree.

3.5.1.2 Terrestrial conditions

Terrestrial atmospheric conditions provide a further demonstration of the possibility that a classical treatment can be in error. Let us consider a reaction coefficient of $\kappa = 10^{-20}\ \mathrm{m^3\ s^{-1}}$ for terrestrial conditions [4]. In order to obtain the appropriate adsorption rate we may use the expression

$$j_A = S\pi a^2 n \left(\frac{8k_B T_g}{\pi m}\right)^{1/2} \tag{3.67}$$

where S is the sticking coefficient of the molecules on the particle surface (between 0 and 1), a is the particle radius, k_B is the Boltzmann constant, T_g is the gas temperature and m is the mass of the reactant molecules. The value $j_A = 13.5\ \mathrm{s^{-1}}$ corresponds to a gas phase species with molecular mass of 3.3×10^{-25} kg (200 g $\mathrm{mol^{-1}}$) at a concentration of 100 ppt in an atmosphere of 1 bar pressure and 300 K of temperature. It is meant to be typical of a small organic trace species, with unit sticking probability.

The evaporation rate λ_A is estimated from the expression

$$\lambda_A = \frac{3}{K_p a \rho}\left(\frac{k_B T}{2\pi m}\right)^{1/2} \tag{3.68}$$

where ρ is the particle mass density and K_p is the partition coefficient [47], which expresses the way gas species are adsorbed onto particles. K_p was taken to be $4 \times 10^{-6}\ \mathrm{m^3\ \mu g^{-1}}$ for a 0.1 μm particle [48], which is again characteristic of a trace, fairly non-adsorptive atmospheric species. With a molecular mass of 3.3×10^{-25} kg, this provides a value $\lambda_A = 333\ \mathrm{s^{-1}}$, corresponding to a residence time of about 3 ms.

Figure 3.12a shows the mean population of A-molecules as a function of its adsorption rate, under the above mentioned atmospheric condition where the aerosol

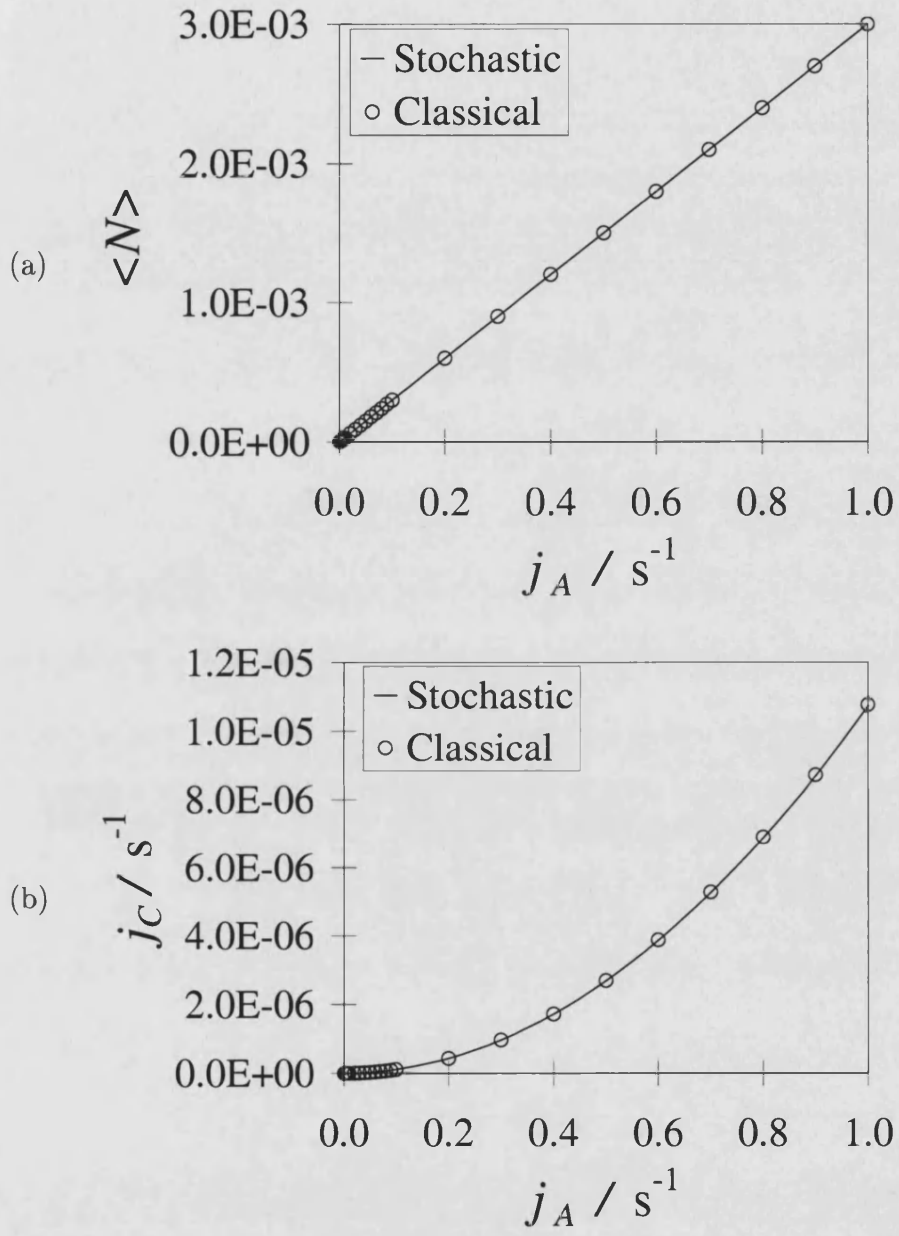

Figure 3.12: Terrestrial conditions with 0.1 μm radius dust particles and $\lambda_A = 333$ s^{-1}. (a) Average number of A-molecules present on the particle surface and (b) dimer production rate (j_C) as a function of j_A. The reason for the lack of difference between the classical and stochastic approaches here is the corresponding high value of $\nu \approx 278$. Any discrepancy would require roughly $\nu < 0$ (see Fig. 3.11).

radius is 0.1μm. The dimer production rate for the same conditions is shown in Figure 3.12b. The mean population ($\bar{N}$) is well below unity in the range of j_A shown and one would have expected the classical approach to differ from the stochastic model in this parameter space. Yet remarkably no distinction appears between the two approaches. The reason for this seemingly peculiar result is that the values of X and ν corresponding to the atmospheric conditions at $j_A = 13.5\ s^{-1}$ are $X \approx 4.05 \times 10^{-2}$ and $\nu \approx 278$. The rather large value of ν leads to the classical limit discussed in §3.3.1.5; for the stochastic model to differ from the classical approach the value of ν must be much smaller.

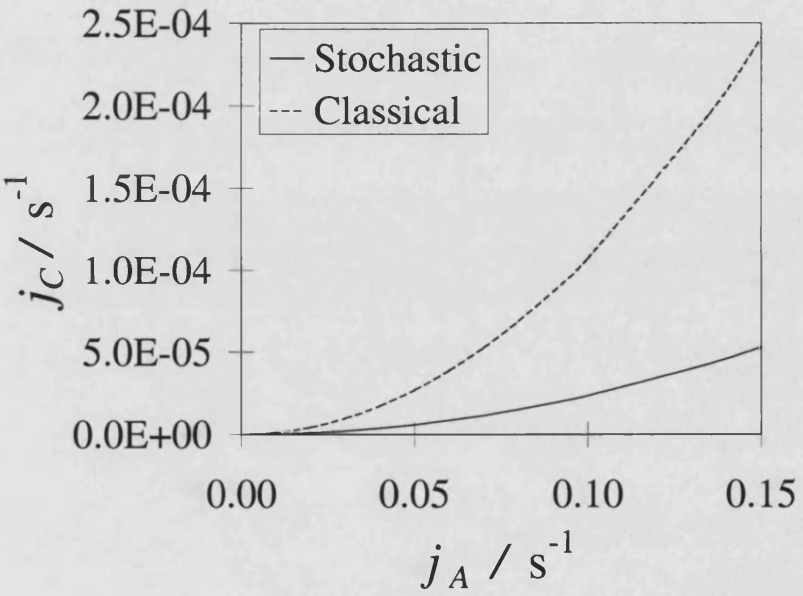

Figure 3.13: Terrestrial conditions with 0.01μm dust particles. Dimer production rate (j_C) as a function of j_A is shown. Corresponding $\nu = -0.72$. At larger values of j_A than shown here, the two curves start to converge and hence not shown here. Chemistry on a 0.01μm particle, however, may not constitute a realistic case.

One way of having a reasonably small value of ν under terrestrial atmospheric conditions is if the particle size is $\sim 0.01\mu$m, with the rest of the parameters unal-

tered. Figure 3.13 shows the dimer production rate under such terrestrial conditions, for which we get $\nu \approx -0.72$ and X still unchanged at 4.05×10^{-2}. The difference between the classical and stochastic results is noticeable now. $j_A = 13.5\ \mathrm{s}^{-1}$ specifically corresponds to the atmospheric conditions mentioned above, for which $j_C^{\text{stochastic}} = 0.4\ \mathrm{s}^{-1}$ whereas $j_C^{\text{classical}} = 1.29\ \mathrm{s}^{-1}$. Hence the classical rate equations overestimate the dimer production rate by over a factor of 3. However this case may not be very realistic since chemistry on 0.01 micron particles is almost irrelevant; they constitute a negligible fraction of surface area and mass within a cloud.

The conditions for failure of a classical approach are clearly delineated by the conditions $X < 1$ and $\nu < 0$ as shown in Figures 3.9 and 3.10. Particular conditions and cases can easily be judged on the basis of these conditions.

3.5.2 $\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C}$

We now turn to the $\mathrm{A}+\mathrm{B} \rightarrow \mathrm{C}$ reaction. Figure 3.14 shows results for the number of molecules A present on the particle surface in equilibrium as a function of j_A while j_B is held at a fixed value. The model reaction chosen here is $\mathrm{O} + \mathrm{N} \rightarrow \mathrm{ON}$, again under interstellar conditions. Here oxygen is treated as the species A and nitrogen is taken as species B. For a cloud at temperature 10 K, Caselli *et al* [40] estimate the values $j_B = 3.9 \times 10^{-6}\ \mathrm{s}^{-1}$, $\kappa/V = 8.7 \times 10^{-5}\ \mathrm{s}^{-1}$ and $\lambda_A \approx \lambda_B \approx 10^{-23}\ \mathrm{s}^{-1}$.

As in the case of reaction $\mathrm{A} + \mathrm{A} \rightarrow \mathrm{C}$, at low flux the stochastic model predicts a higher mean population of reactant molecules compared to the classical result. Interestingly, once j_A falls below the value of j_B, the classical value of $\bar{N}_A$ immediately collapses to almost zero ($\bar{N}_A \sim 10^{-13}$), while the stochastic value sustains larger non-zero values. Note that as in the previous case, a higher mean population will lead to a lower production rate of the dimer.

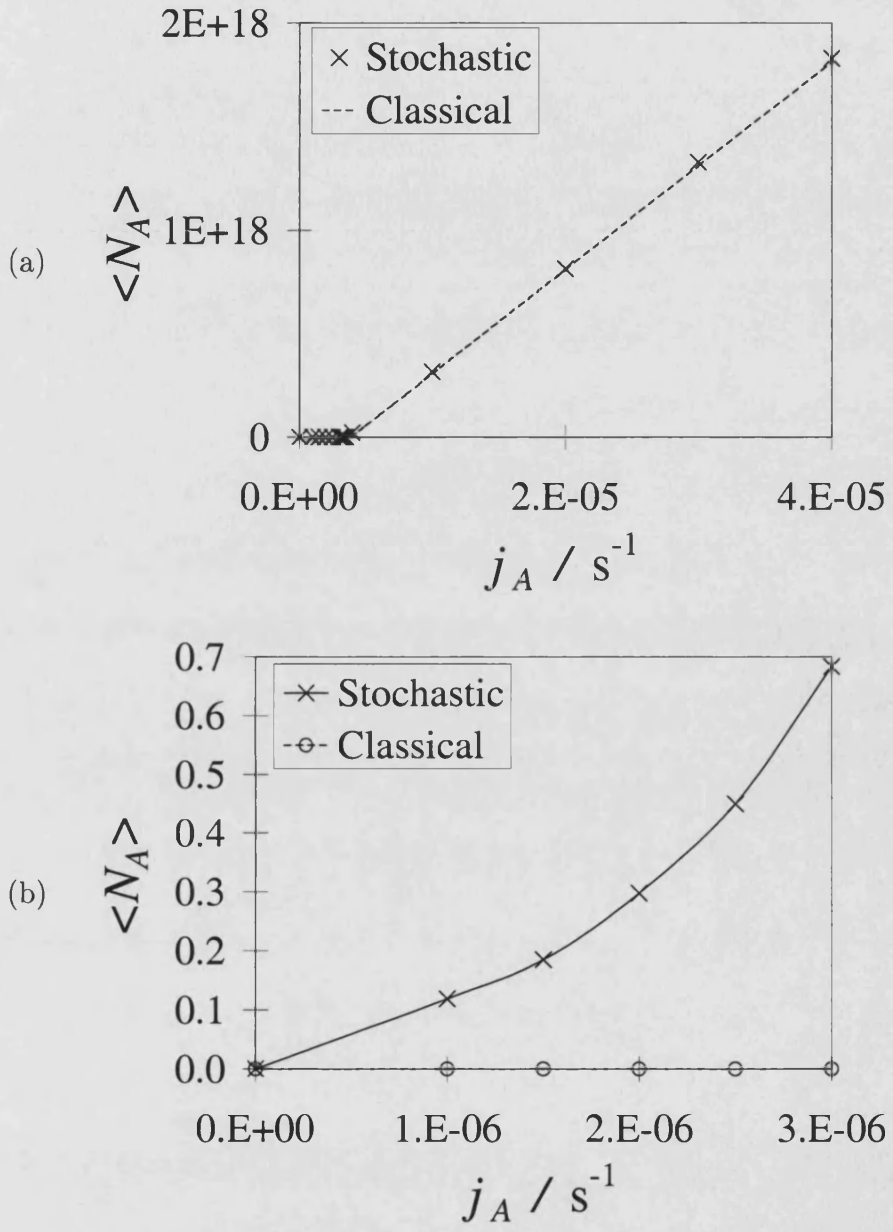

Figure 3.14: In reaction A + B $\rightarrow$ C, average number of molecules A present on the particle surface ($\bar{N}_A$, denoted here as $< N_A >$) in equilibrium as a function of j_A with j_B held at a fixed value. (a) For relatively large range of j_A. (b) At very small flux, the difference between the stochastic and classical models shows up.

Although the numerical techniques developed to solve the kinetics of $A+B \rightarrow C$ are rather difficult to employ under certain conditions and perhaps not always to be recommended, they are nevertheless each valuable in their sphere. Techniques described in the first two methods allow us to gain insight into the likely nature of the analytic solution $F(x, y)$. Whilst the matrix inversion technique of the first method provides the solution in almost no time with the current generation of computers, the second and third methods are also much faster than alternative Monte Carlo approaches. A code developed for the evolution of $W(N, t)$ in the case of reaction $A + A \rightarrow C$, for instance, was found to be about 30 times faster than the Monte Carlo technique for the same reaction as described in §3.4.1. Similarly, it is believed that the evolution of $W(N_A, N_B; t)$ for the $A + B \rightarrow C$ reaction, as explained in the third method, would also be faster than a corresponding Monte Carlo approach.

3.5.3 Possible observational tests

Testing the predictions of this stochastic model against observation is relatively easier in the earth's atmosphere, for obvious reasons, compared to observations of interstellar clouds. In astronomical situations, the only information that can be gained about an object is through electromagnetic radiation coming from the direction of the object. Molecular hydrogen (H_2), for example, does not produce radio emission; it can only be detected through the fact that it absorbs ultraviolet radiation and hence its spectrum will exhibit absorption lines in this region. However, the gas and dust normally becomes so thick in a molecular cloud that the ultraviolet extinction is too large to accurately measure all of the H_2 in the interior of a cloud. In principle though, it should still be possible to estimate the amount of molecular hydrogen. Carbon monoxide emits radio waves due to H_2 molecules colliding with

them, so there is a correlation between the amount of CO and H_2 in the interstellar medium. Hence the easily detected CO radio emission lines can be used to infer the amount of H_2. A decrease in the density of the H_2 gas results in fewer collisions with the CO molecules and a decrease in the CO emission.

Alternatively, a lower production rate of H_2 will mean a higher proportion of atomic hydrogen left in a given interstellar cloud. Hence, it has also been suggested [49] that by observing the easily detectable atomic hydrogen, an estimate of the amount of molecular hydrogen can be made.

3.6 Conclusion

In summary, a stochastic model to describe chemical reactions occurring on the surface of ultra-fine aerosol or interstellar particles under extremely low concentrations of the surrounding reactant gas has been investigated here. Analytic solution is possible in this model for the reaction of type $A + A \rightarrow C$ in the steady state. The solution has been tested against Monte Carlo simulation and the two agree very well. For the reaction of type $A + B \rightarrow C$ analytic solution is not feasible. Hence a number of numerical techniques were developed that seek a solution for the state of the particle surface in terms of the mean population of reactant molecules. The model presented here provides, in general, a time-dependent description of the heterogeneous chemistry mentioned, although the solutions were only obtained for the steady state since the interstellar or atmospheric dust clouds are assumed to have existed for very long. A time-dependent solution to the master equations would be desirable if, for instance, one sought to study industrial processes occurring on much shorter timescales.

As expected, the contrast between the stochastic and classical approaches is most notable at low flux of molecules on the particle surface, although the two approaches agree well when the flux becomes either zero or very large. The result of the simple reaction of type A+A $\rightarrow$ C showed that under interstellar conditions, the classical theory may be overestimating the production rate of molecular hydrogen from its atomic form by over one order of magnitude. Under terrestrial conditions, the overestimation may be lower, but still quite significant. We can conclude that differences in the reaction rate emerge when $X < 1$ and $\nu < 0$, where X and ν are dimensionless combinations of the fundamental process parameters. This is illustrated in Figure 3.11. Similar features emerge for slightly more complex reactions of type A+B $\rightarrow$ C. Especially at very low flux of reactant molecules, when the classical theory predicts almost no reactant population on the dust surface, the stochastic model predicts a larger mean population. Clearly, the use of classical kinetic equations, which are based upon the assumption that the populations of reacting species are large, can be misleading when applied in circumstances where populations are very small.

The stochastic model presented here may be extended to more complex reaction networks by modifying the master equation (3.46), with the variables N_A and N_B replaced by a set of variables $\{N_i\}$, where i would be the number of chemical species participating in the reactions. The exact form of such master equation would be dictated by the way the reactions are interconnected, but this may be done using the same type of arguments as those used in constructing Eqs. (3.2) and (3.46). Such a model is unlikely to have an analytic solution, so the numerical techniques of §3.4.2 will need to be extended accordingly. Computationally this would be a very expensive task in terms of execution time if approaches analogous to the

second and third methods of §3.4.2 are employed. However, the scheme described in the first method essentially relies on matrix inversion, so the computational time consumption is generally tiny once the problem is set up correctly.

Chapter 4

Nucleation: Basic Concepts

In the next two Chapters, we shall consider aspects of the formation and subsequent growth of dust particles. Both of these processes entail the clustering of one or more molecular species, which leads us to consider the phenomenon of nucleation. It is therefore worthwhile sketching a brief description of its basic theoretical concepts [50, 51, 52, 53].

Nucleation is the initial stage of a first order phase transition whereby molecular clusters of a stable phase form out of a metastable phase. It is the first step in many phase transformations and they are very common; dramatic examples can be found in the atmosphere, where the condensation of water vapour, driven below its dew point, gives rise to the formation of water and ice clouds of great variety and beauty [4]. Similar processes on a grander scale are believed to take place in the vicinity of stars, giving rise to equally beautiful dusty nebulae. Even the phase transitions involved in the early Universe soon after the Big Bang are thought to be explicable to a certain extent through the principles of nucleation [7, 54, 55]. Domestic examples are also familiar and processes such as melting, freezing, boiling or condensation are common in industry. An understanding of nucleation is therefore important in both

studying natural phenomena as well as in developing technological applications.

Most of the phase transformations mentioned above are first order, which is to say that a latent heat is transferred during the process and a surface tension exists between the two phases at equilibrium. The transformation usually involves the emergence of assemblages, or clusters, of molecules with characteristics (density, symmetry, etc) of the new phase. However, these clusters are not necessarily all thermodynamically more stable than the original phase. Small clusters, with high proportions of 'surface', tend to be unstable and tend to break apart. For moderate degrees of metastability of the original phase, there exists a 'bottleneck' in the process, corresponding to the need to form a so-called critical molecular cluster. Once one has been formed, further growth is thermodynamically favourable. This is the process of nucleation, driven fundamentally by thermal fluctuations. However, for greater degrees of metastability of the original phase, the phase transformation can become deterministic, with no thermodynamic bottleneck. The process then becomes a second order transformation and proceeds via spinodal decomposition [50].

There are two classes of nucleation, namely homogeneous and heterogeneous. In the former class, which is perhaps more obvious to imagine, monomers of the 'vapour' phase aggregate to form clusters of condensed phase on their own. The latter class involves the need for a host particle that provides a catalytic surface where monomers can gather and form clusters, in a way similar to the heterogeneous chemical reactions described in Chapter 3. Although heterogeneous nucleation is a much more efficient process, the bulk of research has so far been directed towards understanding homogeneous nucleation because of its conceptual simplicity. In outlining the key ideas of nucleation in this Chapter, we keep mainly homogeneous nucleation

in mind, while remembering that the same ideas can normally be extended to the heterogeneous case also.

4.1 Classical rate equations

Nucleation is often treated as a growth-decay ladder of molecular clusters. Consider a vapour cloud full of 'monomers', which could be atoms or even large molecules, that are moving around and occasionally colliding with each other randomly. On collision, two monomers might stick together and then yet another one could stick and form a 'cluster'. The growth of the cluster may thus proceed through attachments of more monomers. The cluster may also decay by loss of monomers, induced, perhaps, by energy input from the surroundings or due to the heat created by the chemical reaction between the monomers. Between these two competing processes, clusters need to reach a critical size i^* before they will, on average, be able to grow further easily. In other words, for clusters consisting of $i < i^*$ molecules, the probability per unit time for a cluster to grow, divided by the probability for it to lose a molecule (decay) is less than unity. For sizes greater than the critical size, the ratio of growth to decay probabilities is greater than unity. Most clusters tend to languish in the sub-critical size region and only occasionally do they manage, by a lucky sequence of growth steps, to reach the critical size and thereafter grow. When a cluster reaches the critical size, nucleation is said to have occurred.

Kinetically, the time-evolution of such a system is modelled using the rate equations

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \beta_{i-1}n_{i-1} - \gamma_i n_i - \beta_i n_i + \gamma_{i+1}n_{i+1} \tag{4.1}$$

for $i \geq 2$, where n_i is the mean population of clusters of size i in the system. β_i is

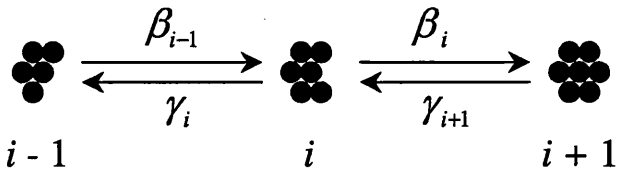

Figure 4.1: A schematic of nucleation kinetics. A cluster of size i may grow or decay due to the addition or removal of a monomer, the rates of these processes being β_i and γ_i respectively.

the rate at which molecules attach themselves to clusters of size i and γ_i is the rate at which molecules are lost from clusters of size i (Figure 4.1). The growth rates β_i are proportional to the number of monomers n_1 in the system, so that we can write

$$\beta_i = \beta_i' n_1. \tag{4.2}$$

The *nucleation rate*, that is the rate at which large clusters are formed, is defined as the difference between $\beta_i n_i$, the number of growth events from size i to $(i+1)$, and $\gamma_{i+1} n_{i+1}$, the number of decays from size $(i+1)$ to i:

$$J = \beta_i n_i - \gamma_{i+1} n_{i+1}. \tag{4.3}$$

By solving the rate equations (4.1) for a steady state situation, Becker and Döring [56] almost 70 years ago arrived at an expression for the nucleation rate,

$$J = \frac{\beta_1 n_1}{1 + \sum_{i=2}^{i_{\max}} \prod_{j=2}^{i} (\gamma_j / \beta_j)}, \tag{4.4}$$

where $i_{\max}$ is the maximum cluster size allowed in the system, which could well be infinity. Equation (4.4) is known as the 'kinetic' expression for the nucleation rate. This solution applies when the growth ladder is terminated by the assumption that clusters at size $i_{\max} + 1$ do not decay, hence $J = \beta_{i_{\max}} n_{i_{\max}}$. For many realistic

situations, the solution is insensitive to the choice of $i_{\max}$, as long as it is large enough. In §5.4.1 we shall explore one way of setting the parameters involved in the above equations. For the moment, this background will be sufficient to proceed to the next Section, and indeed to the next Chapter also.

4.2 Free energy barrier

The same problem can be studied from a different angle by considering the thermodynamic quantities involved. Let us go back to our cloud of vapour and assume it to have a pressure p and temperature T. When this vapour is in thermodynamic equilibrium with a condensate, the vapour is said to be saturated and its chemical potential, μ_v, will be equal to that of the condensed phase, μ_c. The assumption of thermodynamic equilibrium is reasonable since the dust particles are considered to be thermalised through the absorption and emission of radiation in the diffuse ISM (as in Chapter 3), and also through collisions with the surrounding gas molecules in the denser molecular clouds (as in Chapter 6) [15]. A vapour with pressure less than the saturated vapour pressure, $p_{\rm sat}$, is called undersaturated and if $p > p_{\rm sat}$, we have a supersaturated vapour, for which $\mu_v > \mu_c$ (Figure 4.2). More precisely, the *supersaturation*, S, is defined as

$$S = \frac{p}{p_{\rm sat}}, \tag{4.5}$$

where $p_{\rm sat}$ is related to the number density of monomers in the saturated vapour, $\rho_{\rm sat}$, by

$$p_{\rm sat} = \rho_{\rm sat} k_B T \tag{4.6}$$

for an ideal gas.

In reality, a vapour can be kept at pressure $p > p_{\rm new}$ for very long periods with-

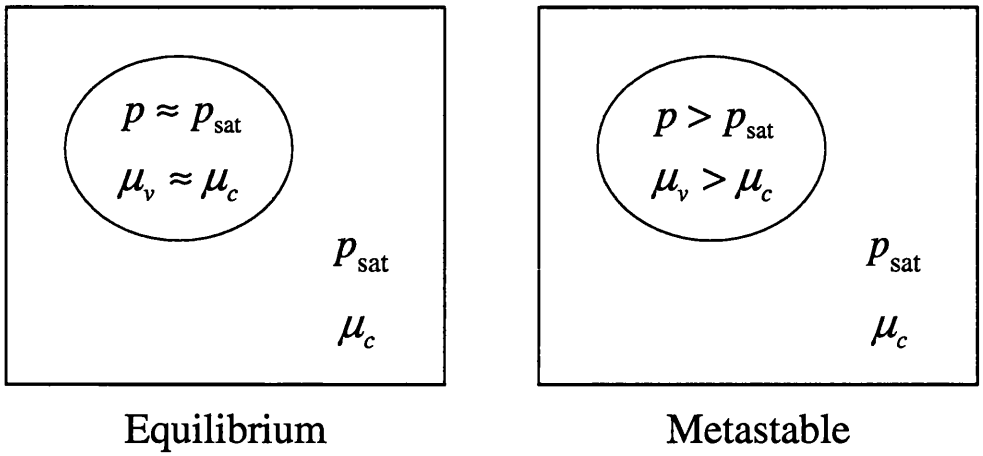

Figure 4.2: Vapour kept in a supersaturated state is metastable with respect to the equilibrium sate.

out observing condensation of the vapour to form the new phase. Although having a supersaturated 'parent' state is a necessary precondition for nucleation (homogeneous nucleation in particular), it does not necessarily occur just with the fulfilment of this condition. Supersaturated old phase can still remain in the metastable state, in quasi-thermodynamic equilibrium for a certain time. Likewise, a liquid can be 'superheated' above its usual boiling temperature, or can be 'supercooled' below its normal freezing temperature. The reason for this is that condensation of subcritical clusters is characterised by an increase in the free energy, $\mathcal{F}$, of the cluster with respect to the pure vapour phase, which is often termed as the *work of formation* of the cluster. The free energy of a cluster of size i is given by

$$\mathcal{F}(i) = \mathcal{E}(i) - T\,\mathcal{S}(i), \tag{4.7}$$

where $\mathcal{E}(i)$ and $\mathcal{S}(i)$ are the total cluster energy and the entropy respectively, and the work of formation can be written as

$$\mathcal{W}(i) = \mathcal{F}(i) - i\mathcal{F}_{\text{mon}}, \tag{4.8}$$

where $\mathcal{F}_{\text{mon}}$ is the free energy of a single monomer in vapour phase. Here we have

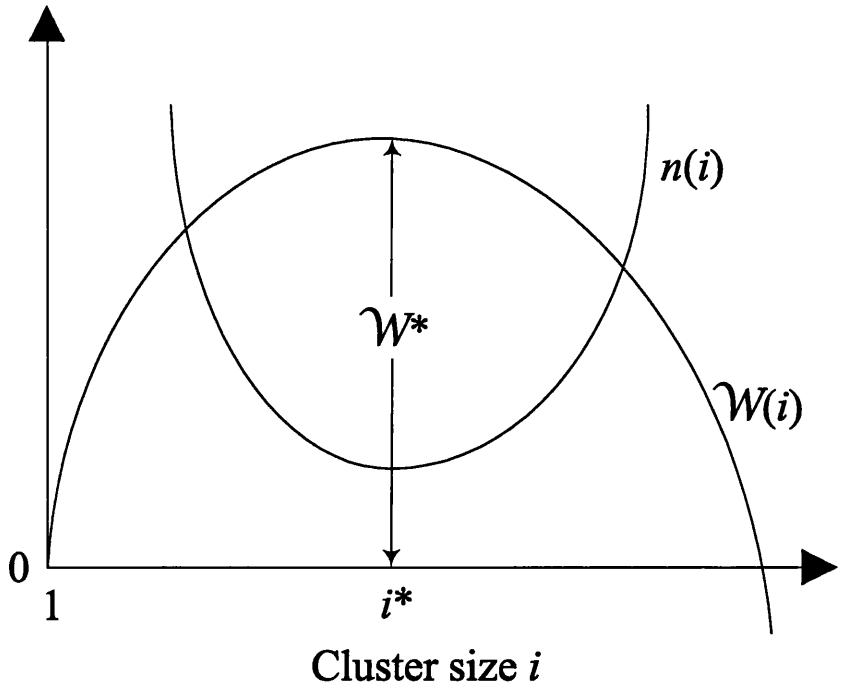

Figure 4.3: The free energy barrier to nucleation. The creation of a cluster of the critical size i^* requires the highest work of formation and clusters of this size are the least probable to be found in the system as the constrained equilibrium distribution $n(i)$ shows a minimum at i^*.

used the Helmholtz free energy, though sometimes the Gibbs potential is also employed. In order for nucleation to occur, the system has to climb a free energy barrier as more and more monomers are attached, before the cluster becomes supercritical, whereupon the system free energy decreases as a function of the cluster size i. It is for this reason that nucleation phenomena can be very slow. Quantitatively, the height of this barrier is equal to the work of formation of the cluster of critical size i^* (Figure 4.3). The constrained equilibrium distribution n_i, which is equal to $\exp(-\mathcal{W}(i)/k_BT)$ [57], shows a minimum at i^*, indicating that the critical size is the most difficult to reach compared to any other sizes in a nucleation process.

The key question then to answer is, what is the nucleation rate for a given

supersaturation and a temperature? The nucleation rate is often written as

$$J = \mathcal{Z}\beta^* \exp\left(-\frac{\mathcal{W}^*}{k_B T}\right), \tag{4.9}$$

where $\mathcal{W}^*$ is the thermodynamic work required to form a critical cluster of the new phase (particle) out of the old phase (vapour). β^* is the rate at which monomers are attached to a cluster of critical size i^*. $\mathcal{Z}$ is the so-called Zeldovich factor and is a dimensionless number, which is given by

$$\mathcal{Z} = \left[\frac{-1}{2\pi k_B T}\frac{\mathrm{d}^2\mathcal{W}(i)}{\mathrm{d}i^2}\right]^{1/2}. \tag{4.10}$$

It accounts for the fact that critical and supercritical clusters can also re-evaporate occasionally due to Brownian motion of the molecules.

The work of formation can be thought to comprise a cost and a payback,

$$\mathcal{W}(i) = \mathcal{W}_{\text{cost}}(i) + \mathcal{W}_{\text{payback}}(i). \tag{4.11}$$

The payback is the change in bulk free energy associated with the conversion of i monomers of vapour into condensed phase. This is simply i times the difference in free energies per monomer between the two phases. Since the free energy per monomer in any bulk phase is just the chemical potential of that phase,

$$\mathcal{W}_{\text{payback}}(i) = -i(\mu_v - \mu_c), \tag{4.12}$$

where μ_v and μ_c are the chemical potentials of vapour and the condensed phase respectively. When the vapour is supersaturated, or equivalently metastable with respect to the condensed phase, its chemical potential is higher and this payback contribution helps to drive the nucleation by reducing $\mathcal{W}(i)$.

Now, for an ideal gas, the chemical potential is related to the vapour pressure p via

$$\mu_v = k_B T \ln(\rho_v \lambda^3), \tag{4.13}$$

where ρ_v is the monomeric number density in the vapour, which is given by

$$\rho_v = p/k_B T. \tag{4.14}$$

λ is the so-called thermal de Broglie wavelength and is expressed as

$$\lambda = \frac{h}{(2\pi\, m\, k_B T)^{1/2}}, \tag{4.15}$$

where m is the mass of a single monomer and h is Planck's constant. At saturated vapour pressure, the chemical potentials of the two phases are by definition equal, namely $\mu_v = \mu_c$, so using Equations (4.13) and (4.6)

$$\mu_c = k_B T \ln(\rho_{\text{sat}} \lambda^3) = k_B T \ln\left(\frac{p_{\text{sat}}}{k_B T}\lambda^3\right). \tag{4.16}$$

Hence, substituting Equations (4.13) and (4.16) into (4.12), one obtains

$$\begin{aligned} \mathcal{W}_{\text{payback}}(i) &= -i\, k_B T \left[\ln\left(\frac{p}{k_B T}\lambda^3\right) - \ln\left(\frac{p_{\text{sat}}}{k_B T}\lambda^3\right)\right] \\ &= -i\, k_B T \ln S, \end{aligned} \tag{4.17}$$

where we have used the definition of the supersaturation, S, from Equation (4.5).

The cost of forming a cluster of the new phase is written in terms of the so-called excess free energy, which is

$$\mathcal{F}_{\text{excess}} = \mathcal{W}_{\text{cost}}(i) = \mathcal{F}(i) - i\mu_c. \tag{4.18}$$

Hence, from Equations (4.11), (4.17) and (4.18), the overall work of formation can be written as

$$\mathcal{W}(i) = [\mathcal{F}(i) - i\mu_c] - i\, k_B T \ln S. \tag{4.19}$$

Using the fact that $k_B T \ln S = \mu_v - \mu_c$ from Equations (4.5), (4.13) and (4.16), this can be simplified as

$$\mathcal{W}(i) = \mathcal{F}(i) - i\mu_v. \tag{4.20}$$

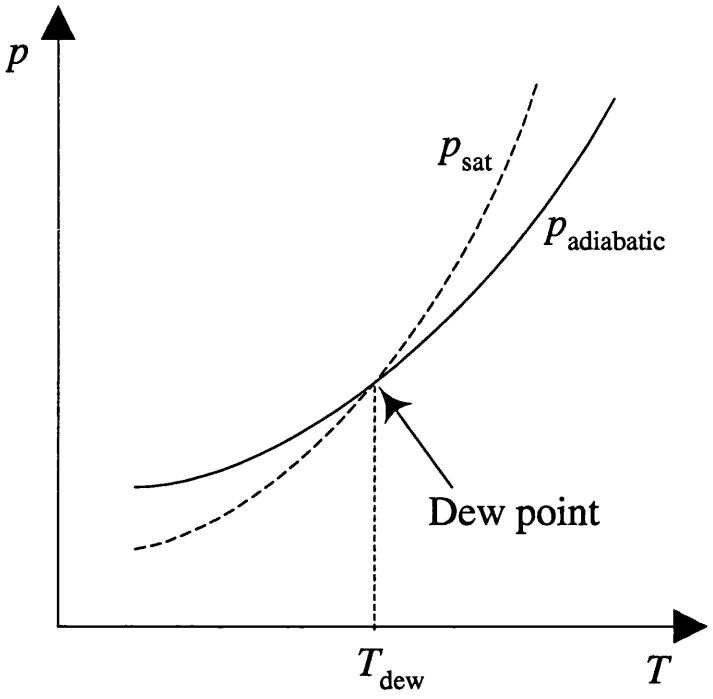

Figure 4.4: Adiabatic cooling of a parcel of vapour. As the vapour is cooled to below the dew point temperature, its pressure becomes greater than the saturated vapour pressure, creating the right condition for nucleation to take place.

The saturated vapour pressure is a function of temperature. If the temperature of a parcel of vapour falls through adiabatic cooling, for example, then the pressure changes, and so does the saturated vapour pressure (Figure 4.4). When the so-called *dew point* is reached, the vapour is said to be at the saturated vapour pressure. If the temperature is dropped further, $\ln S$ will be positive and there is a thermodynamic payback driving the nucleation. This is why one observes nucleation below the dew point temperature.

4.3 Partition function and free energy

Since this treatment of nucleation will be applied to diatomic solid condensation in Chapter 6, let us consider a cluster of such type with stochiometry XY. If the material is solid and the temperature is reasonably high, the atoms would oscillate with high frequency and small amplitude, which enables us to treat the oscillations as harmonic. In this approximation, the free energy appearing in Equation (4.20) for a crystal structure can be expressed as a sum of the temperature-independent potential energy, U_0, associated with the interatomic interactions, and the free energy associated with the normal modes of lattice oscillations [58],

$$\mathcal{F} = U_0 + \mathcal{F}_{\text{modes}}. \tag{4.21}$$

In general, the normal modes arise from vibrational, translational and rotational motion of the atoms and $\mathcal{F}_{\text{modes}}$ can be written in terms of a phase space integral. This is done by remembering that

$$\mathcal{F}_{\text{modes}} = -k_B T \ln Z, \tag{4.22}$$

where the partition function Z is written in the form of the integral

$$Z = \frac{1}{(i!)^2\, h^{6i}} \int \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{p}\, \exp(-\mathcal{H}/k_B T). \tag{4.23}$$

Here, $\mathbf{r}$ and $\mathbf{p}$ are the positions and momenta of the atoms and $\mathcal{H}$ is the Hamiltonian of the system. i is the number of monomers in the cluster and since each monomer is a diatomic molecule, there are $6i$ degrees of freedom in total within the cluster. Therefore, there are $(i!)^2$ number of ways of arranging i indistinguishable atoms of the X type and the same number of indistinguishable atoms of the Y type. This partition function can be factorised into translational, rotational and vibrational parts through a transformation to some suitable relative positions and momenta, so

that

$$Z = Z_{\text{tran}} Z_{\text{rot}} Z_{\text{vib}}, \tag{4.24}$$

and

$$\mathcal{F}_{\text{modes}} = \mathcal{F}_{\text{tran}} + \mathcal{F}_{\text{rot}} + \mathcal{F}_{\text{vib}}. \tag{4.25}$$

We focus on the vibrational degrees of freedom first. The integral can be evaluated in this case by considering the normal modes of vibration obtained from the dynamical matrix of force constants from a molecular dynamics simulation of the system, further details of which we shall discuss in §6.4. Z_{vib} then becomes a product of partition functions for $6i - 6$ oscillators, each of which in the classical limit of high temperatures takes the form $k_B T/\hbar\omega_k$, where ω_k $(k = 1, 2, \ldots, 6i - 6)$ are the angular frequencies of the vibrational modes [59, 60]. Hence the vibrational free energy can be written as

$$\mathcal{F}_{\text{vib}} = -k_B T \ln Z_{\text{vib}} = -k_B T \ln \left(\prod_{k=1}^{6i-6} \frac{k_B T}{\hbar\omega_k} \right) \tag{4.26}$$

There is also an alternative method for obtaining the vibrational modes, which is discussed in Appendix B.

Lifting the restriction of zero angular momentum for the configuration would now allow us to determine the rotational part of the partition function. It has been shown [61] that the result of doing the integration for rotational motion is

$$Z_{\text{rot}} = \frac{\pi^{1/2}}{\chi} \left(\frac{8\pi^2 k_B T}{h^2} \right)^{3/2} (I_1 I_2 I_3)^{1/2}, \tag{4.27}$$

where I_1, I_2 and I_3 are the three principal moments of inertia of the cluster about the centre of mass when the atoms are in their mean positions. χ is the symmetry number associated with a given crystal geometry, which corrects for repeated counting of indistinguishable orientations of the crystal in three dimensions.

Finally, we need the translational free energy of the centre of mass. The partition function for this contribution is

$$Z_{\text{tran}} = V\left(\frac{2\pi(\sum_i m)k_BT}{h^2}\right)^{3/2}, \tag{4.28}$$

where V is the volume of the container in which the cluster is able to translate freely.

Using Equations (4.21), (4.22) and (4.24), together with (4.26), (4.27) and (4.28), we can now calculate the free energy of the cluster,

$$\begin{aligned}\mathcal{F}(i) &= U_0(i) - k_BT\ln(Z_{\text{tran}}Z_{\text{rot}}Z_{\text{vib}}) \\ &= U_0(i) - k_BT\ln\left[V\left(\frac{2\pi(\sum_i m)k_BT}{h^2}\right)^{3/2}\frac{\pi^{1/2}}{\chi}\left(\frac{8\pi^2k_BT}{h^2}\right)^{3/2}(I_1I_2I_3)^{1/2}\right. \\ &\qquad \left.\prod_{k=1}^{6i-6}\frac{k_BT}{\hbar\omega_k}\right],\end{aligned} \tag{4.29}$$

where $U_0(i)$ is the system potential energy at the mean atomic positions.

4.4 Chemical potential

The chemical potential of vapour that appears in Equation (4.20) can be worked out by remembering again that in a bulk phase the chemical potential is equal to the free energy per monomer of that phase. Hence the chemical potential of a single diatomic monomer of vapour is

$$\mu_v = U_0^{\text{mon}} - k_BT\ln\left[\frac{1}{\rho_v}\left(\frac{2\pi\, m\, k_BT}{h^2}\right)^{3/2}\left(\frac{8\pi^2k_BT}{h^2}I_{\text{mon}}\right)\frac{k_BT}{\hbar\omega_{\text{mon}}}\right], \tag{4.30}$$

where U_0^{mon} is the potential energy of a single monomer when the individual atoms within it are at their mean position, ρ_v is the number density of monomers in the vapour, I_{mon} is the moment of inertia of a monomer about its centre of mass and ω_{mon} is the angular frequency of the vibrational mode of the free monomer.

4.5 Cluster densities and nucleation rate

Now that we have worked out a way of evaluating all the quantities required in the implementation of Equation (4.20), the nucleation rate can be evaluated. We calculate the nucleation rate per unit volume, J_V, and assume that the 'container' would be very large in naturally occurring processes. Hence, from Equation (4.9),

$$J_V = \frac{J}{V} = \mathcal{Z}\beta^* \exp\left(-\frac{\{\mathcal{W}^* + k_B T \ln V\}}{k_B T}\right), \tag{4.31}$$

where $\mathcal{W}^*$ is calculated at the critical cluster size i^* with the help of Equations (4.20), (4.29) and (4.30). The term $k_B T \ln V$ would cancel out with a similar term in the expression for $\mathcal{F}(i)$, leaving J_V independent of the volume.

The number of i-clusters, per unit volume is given by

$$\rho_i = \frac{n_i}{V} = \exp\left(-\frac{\{\mathcal{W}(i) + k_B T \ln V\}}{k_B T}\right). \tag{4.32}$$

4.6 Effective work of formation

If one wants to look at the work of formation, $\mathcal{W}(i)$, itself, then the following procedure is useful. Classically, the nucleation rate per unit volume is often written as

$$J_{V,\text{clas}} = \mathcal{Z}\beta^* \, \rho_v \exp\left(-\frac{\mathcal{W}^*_{\text{clas}}}{k_B T}\right), \tag{4.33}$$

where ρ_v is the number density of the vapour monomers and is related to the vapour pressure p by

$$\rho_v \, k_B T = p. \tag{4.34}$$

We can write Equation (4.31) as

$$J_V = \mathcal{Z}\beta^* \, \rho_v \exp\left(-\frac{\{\mathcal{W}^* + k_B T \ln V\}}{k_B T} - \ln \rho_v\right), \tag{4.35}$$

Comparing Equations (4.33) and (4.35) then allows us to write the effective classical work of formation at the critical cluster size i^*, which can then be generalised for all cluster sizes as

$$\mathcal{W}_{\text{eff}}(i) = \mathcal{F}(i) - i\mu_v + k_B T \ln(\rho_v V). \tag{4.36}$$

The term $\mathcal{F}(i) + k_B T \ln V$ would be independent of the volume and knowledge of the number density of the vapour would thus allow us to estimate the work of formation.

Chapter 5

Stochastic Kinetics of Heterogeneous Mantle Growth

5.1 Introduction

We now turn to the problem of the growth of dust particles in the interstellar medium (ISM). In the chemical reactions discussed in Chapter 3, the binding between the molecular species (A and B) and the dust particle was only physical in that it was reasonably easy for the atoms and molecules to break free from the dust surface. However, if these atoms and molecules manage to form chemical bonding with the surface atoms, molecules would accumulate on the surface, enabling formation of a mantle around the core particle. Unlike the newly formed H_2 molecule, which generally detaches from the dust grain after formation through heterogeneous chemistry, heavier molecules like OH, CH and NH would tend to remain attached to the grain and undergo further grain surface reactions with incoming molecules. This process then leads to condensation on the dust particle through the process of heterogeneous nucleation and growth. Incidentally, this growth eventually results

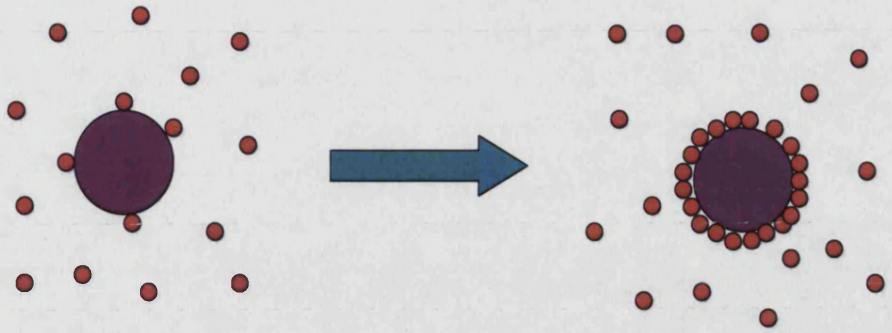

Figure 5.1: Heterogeneous mantle growth occurs through monomers adsorbing one by one onto the dust surface, eventually forming a mantle around the particle.

in the formation of macroscopic pebbles that convert into the building blocks of large heavenly structures like stars, planets and so on. However, there are also more subtle roles these mantles can play in the chemical evolution of the ISM. Icy mantles often act as storage of molecular species. Or their chemical composition could be a suitable substrate for further production of new molecules which would have been almost impossible on the surface of the original grain core or in the pure gas phase.

Most research into nucleation is concerned with the homogeneous process, where the metastability of the original phase is overcome without the presence of special nucleation sites in the system. The critical clusters form in the absence of foreign bodies and container surfaces. However this is not the process responsible for most of the familiar phase transformation examples described in the previous Chapter. The terrestrial atmosphere, for example, is not entirely free of suspended matter and cloud formation takes place by a process of heterogeneous nucleation of water on these particles. The water clusters, and ultimately the cloud droplets, form on the surfaces of suspended particles called cloud condensation nuclei (CCN), since it is far easier thermodynamically to do this than to form a critical cluster homogeneously [62, 63, 64, 65].

Cloud condensation nuclei are solid or liquid aerosols, often only a fraction of a micrometre in diameter. Now, the metastability of a vapour is measured in terms of its supersaturation S, defined as the ratio of the vapour pressure to the saturated vapour pressure, and the critical supersaturation required to drive nucleation at a given rate is a measure of the ease with which critical clusters can be formed. While a value of S of order 10 might be necessary in some circumstances to drive homogeneous nucleation, only $S \sim 0.01$ is sufficient to drive the heterogeneous process if CCN surfaces are present [66, p. 287]. In the atmosphere, supersaturations are usually limited to these values and even smaller values of S are to be found in the low densities of the interstellar medium, so heterogeneous nucleation is the dominant process.

Heterogeneous nucleation has been previously investigated via the free energy calculation approach [67, 68]. However, in understanding the kinetics, a lot more is to be desired for and the rate at which this process occurs is not easy to predict. It is generally considered that the kinetics of nucleation in general were correctly described by Becker and Döring [56] through the solution presented in form of Equation (4.4). Usually, the slightly unrealistic steady state situation is assumed, where the supersaturation of the original phase is held constant in spite of the consumption of material in the formation of new phase. Nevertheless, this is a reasonable approximation when the rate of consumption is low, and so the processes of homogeneous and heterogeneous nucleation are considered to be well represented by formula (4.4) for the nucleation rate.

However, the Becker-Döring approach makes an assumption about the kinetics which may not be valid. The rate equations (4.1) are what we might call *classical* in that the number of growth transitions from size i to $(i+1)$, for example, is taken to

be the population of i-clusters, n_i, multiplied by a rate coefficient β_i proportional to n_1. If n_1 were a precise constant, then this assumption would be valid, but in fact all cluster populations in the problem, including n_1, display fluctuations about a mean value, since the processes of growth and decay occur as stochastic events. As we shall show in the next Section, the growth rate actually requires us to evaluate the mean of the product of the populations of monomers and i-clusters, rather than the product of the mean.

The error incurred by the neglect of fluctuations is small when the populations of clusters are large, by the usual statistical arguments. This is almost always the case in practical cases of homogeneous nucleation: the system is a sample of vapour, say, in a macroscopic container, so that the number of monomers present in the system is huge. However, when the process under consideration is heterogeneous nucleation taking place on the surface of a microscopic particle, such as in the atmosphere or in the interstellar medium, the possibility arises that populations could be small. An experiment involving vapour condensation could be conducted in a macroscopic container, but the actual 'reaction vessel' would be the surface of one of the many particles suspended inside the container. In experiments involving heterogeneous nucleation, therefore, it is possible for the Becker-Döring kinetics to be inappropriate.

It is this possibility that we investigate in this Chapter. There have been some attempts at considering the discrete nature of the nucleating molecules with the aid of stochastic arguments. In particular, Ebeling *et al.* have examined a master equation approach in dealing with the nucleation kinetics [69]. To a limited extent, it is similar to what is proposed in the next Section, but the theory of Ebeling *et al.* gives only a general picture of the kinetics and is not intended for treating

small systems with tiny mean populations of molecules. The possibility of low mean populations encountered in precipitation in small droplets has been considered by Manjunath *et al.*, through stochastic simulations involving a series of the so-called product density equations [70].

Here we consider the complete solution to the heterogeneous nucleation kinetics of growth and decay of clusters of various sizes, where the possibility of fluctuations is properly taken into account [6]. This requires the setting up and solving of master equations for the probability distributions of cluster populations, much the same way as was done in studying dimer formation in Chapter 3. A simple set of rate coefficients is considered, which allows us to perform the computational tasks in an efficient manner and contrast the resulting nucleation rate with the Becker-Döring solution. The conditions necessary for large differences to exist between the 'classical' Becker-Döring solution and the more appropriate 'stochastic' solution to the master equations have been exposed.

5.2 Stochastic approach to the kinetics of heterogeneous nucleation

5.2.1 Rate equations approach

Our system of interest now comprises a host particle surrounded by gas phase molecules (monomers) that occasionally strike and stick to the particle. Once adsorbed, such a monomer may move around the particle during which it may encounter more monomers present on the surface and in due course form a cluster. There is also a possibility that the cluster may decay by loss of monomers.

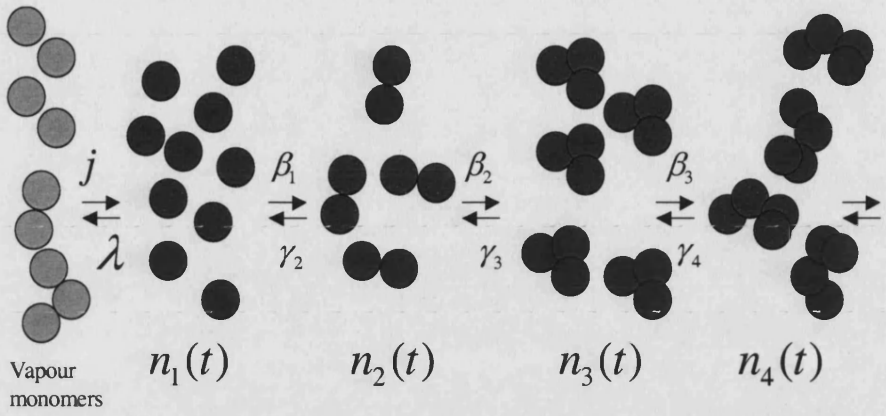

Figure 5.2: Kinetics of heterogeneous nucleation. Here the evolution of the system proceeds via attachment and loss of monomers only, leading to a size distribution n_i at any given time t.

For convenience, let us reproduce Equation (4.1) here, which described the kinetic evolution of the system,

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \beta_{i-1}n_{i-1} - \gamma_i n_i - \beta_i n_i + \gamma_{i+1}n_{i+1} \tag{5.1}$$

for $i \geq 2$, where,

$$\beta_i = \beta_i' n_1. \tag{5.2}$$

In the case of heterogeneous nucleation, the special case for $i = 1$ has to be expressed through a separate equation to accommodate the fact that the 'system' now is only the surface of the particle and monomers ($i = 1$) are being added from *outside* the system, i.e. from the vapour phase. The dynamics for the monomer are expressed by

$$\begin{aligned}\frac{\mathrm{d}n_1}{\mathrm{d}t} &= j - \lambda n_1 - 2\beta_1 n_1 + 2\gamma_2 n_2 - (\beta_2 n_2 - \gamma_3 n_3) - (\beta_3 n_3 - \gamma_4 n_4) - \cdots - \beta_{i_{\max}} n_{i_{\max}} \\ &= j - \lambda n_1 - 2(\beta_1 n_1 - \gamma_2 n_2) - \sum_{i=2}^{i_{\max}-1} (\beta_i n_i - \gamma_{i+1} n_{i+1}) - \beta_{i_{\max}} n_{i_{\max}},\end{aligned} \tag{5.3}$$

where j is the source rate at which monomers attach themselves to the surface from the surrounding medium and λ is the evaporation rate of monomers from the particle

surface. The population of monomers goes down by one when a monomer attaches to a dimer in order to make a trimer, but n_1 increases every time a trimer decays to a dimer, freeing up a monomer. This is reflected in the terms $-(\beta_2 n_2 - \gamma_3 n_3)$. The creation (or decay) of a dimer, however, consumes (or releases) two monomers and this is represented by the terms $-2\beta_1 n_1 + 2\gamma_2 n_2$. Clusters of size larger than $i_{\max}$ are assumed to be stable enough not to decay.

When fluctuations in populations about mean values are taken into account, it would seem reasonable that the rate equations (5.1) should be replaced by something like

$$\frac{\mathrm{d}\langle N_i\rangle}{\mathrm{d}t} = \beta'_{i-1}\langle N_1 N_{i-1}\rangle - \gamma_i\langle N_i\rangle - \beta'_i\langle N_1 N_i\rangle + \gamma_{i+1}\langle N_{i+1}\rangle, \tag{5.4}$$

where the angled brackets represent an averaging over the fluctuations and the cluster populations are written in upper case N_i to remind us that they are fluctuating stochastic variables. Equation (5.3) would similarly be replaced. We shall see in the next Section how such equations can be derived from a stochastic treatment of the populations and how the averages can be evaluated.

5.2.2 Stochastic approach

In the stochastic approach we consider a probability distribution that describes the state of the system in terms of the exact populations of all the allowed cluster sizes. Let the probability that the system contains N_1 monomers, N_2 dimers and in general N_i i-clusters at time t be $W(N_1, N_2, \ldots, N_i, \ldots, N_{i_{\max}}; t) \equiv W(\{N_i\}; t)$. In order to limit the number of elements in this array, we introduce a maximum cluster size $i_{\max}$. We also limit each N_i to be less than or equal to $N_i^{\max}$. The rate of change of this probability is then given by

$$
\begin{aligned}
\frac{\mathrm{d}W}{\mathrm{d}t} &= jW(N_1-1,\ldots) - jW(\ldots) \\
&+ \lambda(N_1+1)W(N_1+1,\ldots) - \lambda N_1 W(\ldots) \\
&+ \beta_1'(N_1+2)(N_1+1)W(N_1+2,N_2-1,\ldots) \\
&- \beta_1' N_1(N_1-1)W(\ldots) \\
&+ \sum_{i=2}^{i_{\max}-1} \beta_i'(N_1+1)(N_i+1)W(N_1+1,\ldots,N_i+1,N_{i+1}-1,\ldots) \\
&+ \beta_{i_{\max}}'(N_1+1)(N_{i_{\max}}+1)W(N_1+1,\ldots,N_{i_{\max}}+1) \\
&- \sum_{i=2}^{i_{\max}} \beta_i' N_1 N_i W(\ldots) \\
&+ \gamma_2(N_2+1)W(N_1-2,N_2+1,\ldots) \\
&+ \sum_{i=3}^{i_{\max}} \gamma_i(N_i+1)W(N_1-1,\ldots,N_{i-1}-1,N_i+1,\ldots) \\
&- \sum_{i=2}^{i_{\max}} \gamma_i N_i W(\ldots). \qquad (5.5)
\end{aligned}
$$

On the right hand side of the above equation, t has been omitted for simplicity. The dots represent values of the N_j that are the same as on the left hand side.

The processes considered are the growth transitions $1+(i-1) \rightarrow i$ and $1+i \rightarrow (i+1)$ due to monomer attachment, as well as the decay processes $i \rightarrow (i-1)+1$ and $(i+1) \rightarrow i+1$ due to monomer detachment from the cluster. The attachment and detachment of dimers, trimers and higher size clusters are neglected. The first two terms (the j terms) describe the addition of a monomer from the surroundings, leading to a monomer population change $N_1 \rightarrow N_1+1$. The third and fourth terms represent loss of a monomer from the particle surface due to the population jump $N_1 \rightarrow N_1-1$. The rest of the terms are constructed using similar arguments for monomeric attachment and detachment to and from dimers, trimers and in general i-clusters. There is a term for $\beta_{i_{\max}}$, but no term involving $\gamma_{i_{\max}+1}$ since clusters at size $i_{\max}$ may grow, but the population at this size receives no additions from the

decay of the next larger cluster. This acts as the boundary condition of the problem.

The classical limit corresponds to the probability distribution W being unity for only one set of possible populations of the i-clusters, that is the mean populations. That is, $W(n_1, n_2, ...n_i...) = 1$ and all other elements are zero. Formally, this is represented, using the Kronecker delta, as

$$W(N_1, N_2, \ldots) = \prod_{i=1}^{i_{\max}} \delta_{N_i n_i} . \tag{5.6}$$

In the steady state and this classical limit, solving equation (5.5) would be equivalent to solving equations (5.1), (5.3) and (4.4), as shown in §5.3.

If Equation (5.5) can be solved by some means, knowledge of W would allow us to generate probability distributions $P_i(N_i)$ for the population of i-clusters:

$$P_i(N_i) = \sum_{(j \neq i)} \sum_{N_j=0}^{N_j^{\max}} W(N_1, \ldots, N_j, \ldots, N_i, \ldots). \tag{5.7}$$

The P_i are likely to look like gaussian distributions for large n_i, or Poisson distributions for small n_i. Ideally, the values of all the $N_j^{\max}$ ought to be infinity for a 'perfect' evaluation of $P_i(N_i)$. However in practice, as we shall see in §5.4.3, satisfactory results may be obtained when the $N_j^{\max}$ are limited to reasonably small values.

It is also possible to calculate joint probabilities, such as $P_{\ell i}(N_\ell, N_i)$, which is the probability that we find N_ℓ ℓ-clusters and N_i i-clusters in the system. These distributions are given by

$$P_{\ell i}(N_\ell, N_i) = \sum_{j \neq \ell, i} \sum_{N_j=0}^{N_j^{\max}} W(N_1, \ldots, N_j, \ldots, N_\ell, \ldots, N_i, \ldots). \tag{5.8}$$

If the steady state elements of W are known, it is possible to calculate the nucleation rate. This is done by summing all the probabilities of growth from any

size i to size $i+1$ and subtract those for decay in the opposite direction:

$$J = \sum_{\{N_j\}} \left(\beta_i' N_1 N_i W(\{N_j\}) - \gamma_{i+1} N_{i+1} W(\{N_j\})\right), \tag{5.9}$$

which by introducing the notation

$$\langle N_i \rangle = \sum_{N_i} N_i P_i(N_i) \tag{5.10}$$

and

$$\langle N_\ell N_i \rangle = \sum_{N_\ell, N_i} N_\ell N_i P_{\ell i}(N_\ell, N_i), \tag{5.11}$$

allows us to write

$$J = \begin{cases} \beta_i' \langle N_1 N_i \rangle - \gamma_{i+1} \langle N_{i+1} \rangle & \text{if } i \geq 2 \\ \beta_i' \langle N_1 (N_i - 1) \rangle - \gamma_{i+1} \langle N_{i+1} \rangle & \text{if } i = 1. \end{cases} \tag{5.12}$$

Any value of $i \geq 2$ in the first of the above expressions would give the same result in the steady state as the nucleation current should be independent of cluster size. If one uses $i = 1$ to compute the nucleation rate, a slight modification is required as in the second expression in Equation (5.12), since having just a single monomer in the system cannot give rise to a nucleation current towards the critical size. In contrast, the nucleation rate given in equation (4.3) according to the standard rate equation (5.1), in the same notation, reads

$$J^{\text{clas}} = \beta_i' \langle N_1 \rangle \langle N_i \rangle - \gamma_{i+1} \langle N_{i+1} \rangle. \tag{5.13}$$

One would expect relative fluctuations in the populations to become negligible when the populations are large, so that a mean of a product becomes the product of the means. It is therefore evident from the comparison of Equations (5.12) and (5.13) that the standard rate equations are valid in the large population limit. It is also possible to visualise how the standard result for the nucleation rate must be modified for small systems. By writing

$$\beta_i' \langle N_1 N_i \rangle = (1 + \epsilon_i) \beta_i' \langle N_1 \rangle \langle N_i \rangle, \tag{5.14}$$

the expression for the rate given in equation (4.4) can be used to see that the nucleation rate, J_{large}, appropriate for a large system may be related, to a good approximation, to the nucleation rate for a small system, J_{small}, by introducing a modification factor ϵ_i such that

$$J_{\text{small}} = J_{\text{large}} \prod_{i=1}^{i^*} (1 + \epsilon_i) \,, \tag{5.15}$$

where i^* is the critical size, where the rate coefficients for growth and decay are equal ($\beta_i = \gamma_i$). We are interested in calculating the modification factor, or equivalently, $J_{\text{small}}/J_{\text{large}}$.

5.3 Reduction of master equations to rate equations

It is possible to show that the master equations (5.5) of heterogeneous nucleation do indeed reduce to the rate equations (5.1) and (5.3) in the classical limit of relatively large populations of clusters. In this Section, we go through the detailed derivation.

To begin with, we define an operator $\hat{O}$ such that

$$\hat{O} \cdot f = \sum_{\{N_i\}=0}^{\infty} N_\ell \cdot f, \tag{5.16}$$

i.e., we multiply the given term f by N_ℓ (where $\ell = 1, \ldots, i_{\max}$) and sum the result over all the $\{N_i\}$. Let us perform this operation on both sides of Eq. (5.5). This makes the left hand side read as

$$\sum_{\{N_i\}=0}^{\infty} N_\ell \frac{\mathrm{d}W(\{N_i\})}{\mathrm{d}t} = \frac{\mathrm{d}\langle N_\ell \rangle}{\mathrm{d}t}, \tag{5.17}$$

which is equivalent to the L.H.S. of Equations (5.1) and (5.3). Now consider the consequence of this operation on the right hand side of Eq. (5.5). On the R.H.S.,

one needs to treat separately the cases of $\ell = 1$ and $\ell > 1$ since there are different rate equations for the two cases of N_1 and N_ℓ ($\ell > 1$) in the classical picture. Let us consider terms proportional to the parameters j, λ, β_i' and γ_i one by one and try to compare them with those found in the rate equations (5.1) and (5.3).

5.3.1 Adsorption term

a. $\ell = 1$

Operating upon the first term in Eq. (5.5) by $\hat{O}$ along with $\ell = 1$ will produce

$$\sum_{\{N_i\}=0}^{\infty} jN_1\, W(N_1 - 1, \ldots).$$

In order to bring the probability W in the same form as on the left hand side, that is $W(N_1, N_2, \ldots, N_{i_{\max}})$, we can make the substitution $N_1 \to N_1 + 1$, which is what happens to the monomer population due to the j term. The above notation will then turn into

$$\sum_{\{N_i\}=0}^{\infty} j(N_1 + 1)W(N_1, \ldots) = j\langle N_1 + 1\rangle. \tag{5.18}$$

The sum over this new N_1 label should run from -1 to ∞, but in that case the unphysical first term in the series would vanish, so the lower limit is indeed zero.

Operating upon the second term in Eq. (5.5) with $\hat{O}$ will give

$$-\sum_{\{N_i\}=0}^{\infty} jN_1\, W(\ldots) \;=\; -j\langle N_1\rangle. \tag{5.19}$$

In the classical limit, the upper case N_1 together with angled brackets is replaced by n_1, so from Equations (5.18) and (5.19), the net result of applying $\hat{O}$ on both the j terms in Eq. (5.5) is

$$j(n_1 + 1) - jn_1 = j. \tag{5.20}$$

This is precisely what we have as the 'j term' in the rate equation (5.3), which was written down explicitly for the monomeric ($\ell = 1$) population.

b. $\ell > 1$

If ℓ is not equal to 1, then the operation due to $\hat{O}$ will make the first term in Eq. (5.5) read

$$\sum_{\{N_i\}=0}^{\infty} jN_\ell W(N_1 - 1, \ldots),$$

where $\ell \neq 1$. This time the substitution $N_1 \rightarrow N_1 + 1$ will lead to

$$\sum_{\{N_i\}=0}^{\infty} jN_\ell W(N_1, \ldots) = j\langle N_\ell \rangle. \tag{5.21}$$

The second term of the master equation under the operation of $\hat{O}$ will be similar to the expression (5.19):

$$-\sum_{\{N_i\}=0}^{\infty} jN_\ell W(\ldots) = -j\langle N_\ell \rangle. \tag{5.22}$$

Hence the sum of Equations (5.21) and (5.22) will be zero and indeed, there is no j term in the rate equations (5.1).

5.3.2 Desorption term

a. $\ell = 1$

If we apply the operator $\hat{O}$ to the third term of the master equation (5.5), we have

$$\sum_{\{N_i\}=0}^{\infty} \lambda N_1(N_1 + 1)\, W(N_1 + 1, \ldots).$$

This time we make the substitution $N_1 \rightarrow N_1 - 1$ so that the above expression is converted into

$$\sum_{\{N_i\}=0}^{\infty} \lambda(N_1 - 1)N_1\, W(N_1, \ldots) = \lambda\langle (N_1 - 1)N_1 \rangle. \tag{5.23}$$

The lower limit for the sum over the shifted variable N_1 should be $+1$, but we can extend this to zero without changing the result of the summation.

Performing the operation $\widehat{O}$ on the fourth term of the master equation will give us

$$-\sum_{\{N_i\}=0}^{\infty} \lambda N_1 N_1 \, W(\ldots) = -\lambda \langle N_1^2 \rangle. \tag{5.24}$$

Once again, to see the correspondence with the classical model, we replace the angled brackets and the upper case N_1 with the lower case n_1, so we are left with the net result

$$\lambda(n_1^2 - n_1) - \lambda n_1^2 = -\lambda n_1. \tag{5.25}$$

This is the λ term found in the monomeric rate equation (5.3).

b. $\ell > 1$

The third term of Eq. (5.5) under the influence of $\widehat{O}$ will this time become

$$\sum_{\{N_i\}=0}^{\infty} \lambda N_\ell (N_1 + 1) \, W(N_1 + 1, \ldots),$$

and the substitution $N_1 \to N_1 - 1$ will make it

$$\sum_{\{N_i\}=0}^{\infty} \lambda N_\ell N_1 \, W(N_1, \ldots) = \lambda \langle N_\ell N_1 \rangle. \tag{5.26}$$

The operation due to $\widehat{O}$ on the fourth term of Eq. (5.5) will give us $-\lambda \langle N_\ell N_1 \rangle$. Hence the lambda term will vanish for the $\ell > 1$ case and is absent in the rate equation (5.1) also.

5.3.3 Growth term

a. $\ell = 1$

If we operate on the fifth term in the master equation (5.5) with $\widehat{O}$, using $\ell = 1$, we get

$$\sum_{\{N_i\}=0}^{\infty} \beta_1' N_1(N_1+2)(N_1+1)\, W(N_1+2, N_2-1, \ldots),$$

which with substitutions $N_1 \to N_1 - 2$ and $N_2 \to N_2 + 1$ becomes

$$\sum_{\{N_i\}=0}^{\infty} \beta_1'(N_1-2)(N_1-1)N_1\, W(N_1, \ldots) = \beta_1\langle (N_1-2)(N_1-1)\rangle. \tag{5.27}$$

where we have used the fact that $\beta_i = \beta_i' N_1$. The sixth term can be operated on without having to do any re-labelling of N:

$$-\sum_{\{N_i\}=0}^{\infty} \beta_1' N_1 N_1(N_1-1)\, W(\ldots) = -\beta_1\langle N_1(N_1-1)\rangle. \tag{5.28}$$

The seventh term will, however, require re-labelling in order to bring W in the desired form. We first operate on it with $\widehat{O}$ to get

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}-1} \beta_i' N_1(N_1+1)(N_i+1)\, W(N_1+1, \ldots, N_i+1, N_{i+1}-1, \ldots),$$

and then use the substitutions $N_1 \to N_1 - 1$, $N_i \to N_i - 1$ and $N_{i+1} \to N_{i+1} + 1$ in order to obtain

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}-1} \beta_i'(N_1-1)N_1 N_i W(N_1, \ldots, N_i, N_{i+1}, \ldots) = \sum_{i=2}^{i_{\max}-1} \beta_i\langle (N_1-1)N_i\rangle. \tag{5.29}$$

With some thought, it is possible to realise that result (5.29) will hold true for any value of i in the $\sum_{i=2}^{i_{\max}-1}$ series. A similar procedure on the eighth term of Eq. (5.5) will give us

$$\beta_{i_{\max}}\langle (N_1-1)N_{i_{\max}}\rangle, \tag{5.30}$$

which essentially completes the series in Equation (5.29) from $i = 2$ to $i_{\max}$. Finally, we operate on the ninth term with $\hat{O}$ and obtain

$$-\sum_{\{N_i\}=0}^{\infty}\sum_{i=2}^{i_{\max}}\beta_i' N_1 N_1 N_i\, W(\ldots) = -\sum_{i=2}^{i_{\max}}\beta_i\langle N_1 N_i\rangle, \tag{5.31}$$

which again holds no matter what value of i is chosen in the $\sum_{i=2}^{i_{\max}}$ series.

Hence the sum of all the β_i' terms in Equations (5.27), (5.28), (5.29), (5.30) and (5.31) will be

$$\sum_{i=2}^{i_{\max}}\beta_i\langle(N_1-1)N_i\rangle - \sum_{i=2}^{i_{\max}}\beta_i\langle N_1 N_i\rangle + \beta_1\Big[\langle(N_1-2)(N_1-1)\rangle - \langle(N_1-1)N_1\rangle\Big]. \tag{5.32}$$

If we now replace the upper case N with its lower case counterpart, discarding the angled brackets to reflect the classical limit, expression (5.32) is easily reduced to

$$-\sum_{i=2}^{i_{\max}}\beta_i n_i - 2\beta_1(n_1-1). \tag{5.33}$$

It can be seen that these are the β terms in the rate equation (5.3) provided that $n_1 - 1 \approx n_1$ in the above expression. This is a fair approximation in the classical limit where the monomeric population is high.

b. $\ell > 1$

Additional care is required when one deals with the case of $\ell \neq 1$ in the β_i' terms. This is due to the series $\sum_{i=2}^{i_{\max}}$ involved and unlike the $\ell = 1$ case, contributions due to different values of i need to be examined explicitly.

Consider the fifth term in Eq. (5.5) first. With the operator $\hat{O}$ applied, it will read

$$\sum_{\{N_i\}=0}^{\infty}\beta_1' N_\ell(N_1+2)(N_1+1)\, W(N_1+2, N_2-1, \ldots),$$

and the substitutions $N_1 \to N_1 - 2$ and $N_2 \to N_2 + 1$ will make it

$$\sum_{\{N_i\}=0}^{\infty} \beta_1' N_\ell N_1(N_1-1)\, W(\ldots) = \beta_1 \langle (N_1-1)N_\ell \rangle \quad \text{if} \quad \ell \geq 3 \tag{5.34}$$

and

$$\sum_{\{N_i\}=0}^{\infty} \beta_1' N_1(N_1-1)(N_2+1)\, W(\ldots) = \beta_1 \langle (N_1-1)(N_2+1) \rangle \quad \text{if} \quad \ell = 2. \tag{5.35}$$

The sixth term in Eq. (5.5) will not require any re-labelling of N after being operated on by $\hat{O}$ and regardless of the value of ℓ it will become

$$-\sum_{\{N_i\}=0}^{\infty} \beta_1' N_\ell N_1(N_1-1)\, W(\ldots) = -\beta_1 \langle (N_1-1)N_\ell \rangle. \tag{5.36}$$

Hence for $\ell \geq 3$ the sum of the positive and negative β_1 terms, given in expression (5.34) and (5.36), is zero. The rate equation (5.1) written down for $i \geq 3$ will surely have no β_1 terms. For the special case of $\ell = 2$, the sum of expressions (5.35) and (5.36) will leave $\beta_1(n_1 - 1)$ in the classical language. Considering the rate equation (5.1) for $i = 2$ case, one would find the term $\beta_1 n_1$, which is approximately equal to the stochastic result $\beta_1(n_1 - 1)$, provided that $n_1 \gg 1$. This is a valid assumption in the classical limit, and so the β_1' terms in the stochastic master equation are reducible to those in the classical rate equations when the mean populations are large.

Let us now consider the seventh term in Eq. (5.5). With operator $\hat{O}$ acting on it, it would read

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}-1} \beta_i' N_\ell (N_1+1)(N_i+1)\, W(N_1+1, \ldots, N_i+1, N_{i+1}-1, \ldots).$$

Consider the expansion of the second summation here:

$$\sum_{\{N_i\}=0}^{\infty} \beta_2' N_\ell (N_1+1)(N_2+1)\, W(N_1+1, N_2+1, N_3-1, \ldots)$$
$$+ \sum_{\{N_i\}=0}^{\infty} \beta_3' N_\ell (N_1+1)(N_3+1)\, W(N_1+1, \ldots, N_3+1, N_4-1, \ldots)$$
$$+ \sum_{\{N_i\}=0}^{\infty} \beta_4' N_\ell (N_1+1)(N_4+1)\, W(N_1+1, \ldots, N_4+1, N_5-1, \ldots) + \cdots$$

With appropriate re-labelling, as done before, and remembering that $\beta_i = \beta_i' N_1$, it is possible to show that this reduces to

$$\beta_{\ell-1}\langle N_{\ell-1}(N_\ell+1)\rangle + \beta_\ell \langle N_\ell (N_\ell - 1)\rangle + \sum_{\substack{i=2 \\ i\neq \ell-1,\,\ell}}^{i_{\max}} \beta_i \langle N_i N_\ell \rangle. \tag{5.37}$$

The operator $\hat{O}$ will reduce the eighth term of Eq. (5.5) into

$$\sum_{\{N_i\}=0}^{\infty} \beta'_{i_{\max}} N_\ell (N_1+1)(N_{i_{\max}}+1)\, W(N_1+1, \ldots, N_{i_{\max}}+1),$$

and with the re-labelling $N_1 \to N_1 - 1$ and $N_{i_{\max}} \to N_{i_{\max}} - 1$ will give us

$$\sum_{\{N_i\}=0}^{\infty} \beta'_{i_{\max}} N_\ell N_1 N_{i_{\max}}\, W(\ldots) = \beta_{i_{\max}} \langle N_{i_{\max}} N_\ell \rangle \quad \text{if} \quad \ell = 2, \ldots, i_{\max} - 1$$
$$\sum_{\{N_i\}=0}^{\infty} \beta'_{i_{\max}} (N_{i_{\max}} - 1) N_1 N_{i_{\max}}\, W(\ldots) = \beta_{i_{\max}} \langle N_{i_{\max}} (N_{i_{\max}} - 1)\rangle \quad \text{if} \quad \ell = i_{\max}. \tag{5.38}$$

A similar argument applies to the ninth term of Eq. (5.5). The operator $\hat{O}$ will reduce this term to

$$-\sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}} \beta_i' N_\ell N_1 N_i\, W(\ldots) = -\sum_{i=2}^{i_{\max}} \beta_i \langle N_i N_\ell \rangle \tag{5.39}$$

regardless of the value of ℓ. Hence summing the seventh, eighth and ninth terms of the master equation, given here as expressions (5.37), (5.38) and (5.39), and replacing the upper case N with the lower case n in the classical picture will give us $\beta_{\ell-1} n_{\ell-1} - \beta_\ell n_\ell$, where $\ell = 2, \ldots, i_{\max}$. These are the β_i terms in the classical rate equation (5.1).

5.3.4 Decay term

a. $\ell = 1$

The effect of the operator $\widehat{O}$, with $\ell = 1$, on the tenth term in the master equation (5.5) will be

$$\sum_{\{N_i\}=0}^{\infty} \gamma_2 N_1(N_2+1)W(N_1-2, N_2+1, \ldots),$$

and the re-labelling $N_1 \to N_1 + 2$ and $N_2 \to N_2 - 1$ will give us

$$\sum_{\{N_i\}=0}^{\infty} \gamma_2(N_1+2)N_2\, W(\ldots) \;=\; \gamma_2\langle(N_1+2)N_2\rangle. \tag{5.40}$$

The eleventh term of Equation (5.5), under the operation due to $\widehat{O}$ will become

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=3}^{i_{\max}} \gamma_i\, N_1(N_i+1)W(N_1-1, \ldots, N_{i-1}-1, N_i+1, \ldots),$$

which with the re-labelling $N_1 \to N_1 + 1$, $N_{i-1} \to N_{i-1} + 1$, $N_i \to N_i - 1$ becomes

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=3}^{i_{\max}} \gamma_i(N_1+1)N_i\, W(\ldots) \;=\; \sum_{i=3}^{i_{\max}} \gamma_i\langle(N_1+1)N_i\rangle. \tag{5.41}$$

The last term in Equation (5.5) is more straight forward and does not require any re-labelling, so the operator $\widehat{O}$ will make it

$$-\sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}} \gamma_i\, N_1 N_i\, W(\ldots) \;=\; -\sum_{i=2}^{i_{\max}} \gamma_i\langle N_1 N_i\rangle. \tag{5.42}$$

Replacing the angled brackets and the upper case N with the lower case n in the classical limit, the sum of all the γ_i terms expressed in (5.40), (5.41) and (5.42) will be

$$2\,\gamma_2 n_2 + \sum_{i=3}^{i_{\max}} \gamma_i\, n_i \;=\; 2\,\gamma_2 n_2 + \sum_{i=2}^{i_{\max}-1} \gamma_{i+1} n_{i+1}. \tag{5.43}$$

These are precisely the γ_i terms appearing in the monomeric rate equation (5.3).

b. $\ell > 1$

If we operate on the tenth term of Equation (5.5) with $\hat{O}$, we get

$$\sum_{\{N_i\}=0}^{\infty} \gamma_2 N_\ell (N_2 + 1) W(N_1 - 2, N_2 + 1, \ldots),$$

and with the re-labelling $N_1 \to N_1 + 2$ and $N_2 \to N_2 - 1$ it becomes

$$\sum_{\{N_i\}=0}^{\infty} \gamma_2 N_\ell N_2 \, W(\ldots) = \gamma_2 \langle N_\ell N_2 \rangle. \tag{5.44}$$

If operated upon by $\hat{O}$, the eleventh term of Equation (5.5) will read

$$\sum_{\{N_i\}=0}^{\infty} \sum_{i=3}^{i_{\max}} \gamma_i N_\ell (N_i + 1) W(N_1 - 1, \ldots, N_{i-1} - 1, N_i + 1, \ldots).$$

With suitable substitutions, it can be shown that this expression is equivalent to

$$\gamma_\ell \langle N_\ell (N_\ell - 1) + \gamma_{\ell+1} \langle N_{\ell+1} (N_\ell + 1) \rangle + \sum_{\substack{i=3 \\ i \neq \ell, \ell+1}}^{i_{\max}} \gamma_i \langle N_\ell N_i \rangle. \tag{5.45}$$

Finally, the last term in Equation (5.5) under the operation due to $\hat{O}$ will appear as

$$- \sum_{\{N_i\}=0}^{\infty} \sum_{i=2}^{i_{\max}} \gamma_i N_\ell N_i \, W(\ldots) = - \sum_{i=2}^{i_{\max}} \gamma_i \langle N_\ell N_i \rangle, \tag{5.46}$$

where $\ell = 2, \ldots, i_{\max}$. Hence the sum of the all the γ_i terms given in (5.44), (5.45) and (5.46) will be

$$\gamma_\ell \langle N_\ell (N_\ell - 1) \rangle + \gamma_{\ell+1} \langle N_{\ell+1} (N_\ell + 1) \rangle - \gamma_\ell \langle N_\ell N_\ell \rangle - \gamma_{\ell+1} \langle N_{\ell+1} N_\ell \rangle,$$

which under the classical limit can be simplified as $\gamma_{\ell+1} n_{\ell+1} - \gamma_\ell n_\ell$. These are the γ_i terms found in the classical rate equation (5.1), except that here the subscript ℓ is used for the labelling purpose.

We therefore conclude that the set of stochastic master equations (5.5) are reducible to the set of classical rate equations given in (5.1) and (5.3) when the mean

populations are large. Furthermore, it is possible to justify the stochastic expression for the nucleation rate, given in Equation (5.12).

Having tested the validity of the master equations in the classical limit, we can now proceed to numerical calculations.

5.4 Model calculations

5.4.1 Parameterisation

The master equations (5.5) are driven by the input parameters j, λ, β_i' and γ_i. In order to investigate the problem of heterogeneous nucleation in small systems, we must carefully choose the input parameters that are likely to lead to small cluster populations.

Let us introduce a size parameter ξ, which may be taken to be proportional to the surface area of the host particle. The coefficients λ and γ_i are the decay rates of monomers ($i = 1$) and i-mers ($i \geq 2$) respectively and hence may be taken as independent of the system size. The attachment rate j of monomers onto the particle surface, however, should increase linearly with ξ. It is useful to consider temporarily the dynamics in the absence of any dimer production, in which case the mean monomer population would be given by a balance between j and λ, namely, $\langle N_1 \rangle \simeq j/\lambda$. If j_0 is the value of j at $\xi = 1$ then we can write

$$j = \xi j_0, \tag{5.47}$$

so that a rough solution to Equation (5.3) is

$$\langle N_1 \rangle \simeq \frac{j_0}{\lambda} \xi. \tag{5.48}$$

For convenience, let us postulate that $\xi = 1$ is the system with a nominal mean monomer population of unity. This imposes the condition $j_0 = \lambda$.

For simplicity, we assume the growth rate β_i' to be independent of the cluster size i, i.e., $\beta_1' = \beta_2' = \cdots = \beta_{i_{\max}}'$. On the other hand, β_i' will be inversely proportional to ξ since it measures the likelihood that an adsorbed monomer will encounter an adsorbed i-mer. As the system gets bigger, this likelihood would diminish. Furthermore, we may fix β_i' such that at $\xi = 1$ the mean growth rate of an i-mer is unity. Remembering from Eq. (5.2) that $\beta_i = \beta_i'\langle N_1 \rangle$, this means that $\beta_i' = 1$ at $\xi = 1$ and in general therefore,

$$\beta_i' = \frac{1}{\xi} \,. \tag{5.49}$$

The choice of the parameters γ_i must satisfy the requirement that at the critical size i^*, a cluster is as likely to decay as it is likely to grow, i.e., $\gamma_{i^*} = \beta_{i^*}'\langle N_1 \rangle$. With the above stated choice of β_i' and $\langle N_1 \rangle$, this means that $\gamma_{i^*} = 1$ at $\xi = 1$. Indeed, this should be true for any value of ξ as the decay rates are independent of the system size. The i-dependence of γ_i may be chosen to express the fact that small clusters are more likely to decay than large clusters. We therefore choose

$$\gamma_i = \left(\frac{i^*}{i}\right)^{\phi} , \tag{5.50}$$

where ϕ is some constant to be decided. Entirely for computational convenience, and without suggesting that the model should represent a real system, we shall choose $\phi = 2$ and $i^* = 2$. This form of γ_i ensures that a cluster below the critical size ($i < i^*$) has a high probability of decay, whereas those above the critical size ($i > i^*$) will find it easier to grow.

The relative values of j_0 (and λ) and β_i' control the degree to which the mean monomer population is close to the estimate (5.48). We shall explore cases where

$j_0 \gg 1$ and $j_0 = 1$ in §5.4.3.

5.4.2 Classical solution

The most convenient way of deducing the classical nucleation rate for a given set of parameters j_0, λ, ξ, β_i and γ_i is through the expression (4.4). However the n_1 appearing in that equation still needs to be known. Although in the large j_0 limit expression (5.48) for $\langle N_1 \rangle$ may provide a reasonable estimate, this is not guaranteed to be true in general. A better method of finding n_1 is as follows.

Equation (5.3) in the steady state may be written, with the help of Eq. (4.3), as

$$\begin{aligned} 0 &= j_0\xi - \lambda n_1 - 2J - (i_{\max} - 2)J - J \\ &= j_0\xi - \lambda n_1 - (i_{\max} + 1)J, \end{aligned} \tag{5.51}$$

where J is given by Eq. (4.4). Let us assign a function

$$g(n_1) = j_0\xi - \lambda n_1 - (i_{\max} + 1)J. \tag{5.52}$$

This function decreases with increasing n_1. As an initial approximation, we provide $n_1 = j_0\xi/\lambda$, which in all practical cases is at least a slight overestimation of the actual value of n_1. We then iteratively search for a zero of the function $g(n_1)$ by subtracting a very small amount (typically $\sim 10^{-6}$) from the trial value of n_1 and evaluating a new value of $g(n_1)$. This process is continued until a solution is found within a very small tolerance. The final value of n_1 that corresponds to $g(n_1) = 0$ can then be utilised in Eq. (4.4) to find the classical value of the nucleation rate.

5.4.3 Solving the master equation

Given all the necessary parameters given in §5.4.1, we are in a position to solve the master equations (5.5) which should ultimately yield the stochastic solution to the system. Solving Eq. (5.5) analytically does not appear to be a feasible task. We therefore look for an appropriate numerical technique to act as a substitute.

Computationally, we discretise time t and replace the $\mathrm{d}t$ by a very small but finite Δt in Equation (5.5). The $\mathrm{d}W(t)$ may then be replaced by $W(t+\Delta t)-W(t)$, thus allowing Eq. (5.5) to be solved iteratively. As an initial condition, we set $W(0,0,0,\ldots,0;t=0)=1$ with all the remaining elements of the array $W(\{N_i\})$ set to zero, specifying an empty system to start with. The system thereafter evolves in time until a steady state is reached.

Equations (5.5) represent a set of coupled differential equations. $i_{\max}$ is the largest size of cluster that can form on the particle and needs to be specified explicitly at the beginning. In principle, it should be large enough so that the contribution due to terms with $i_{\max}+1$ in the series appearing in Equation (4.4) is negligible.

Strictly speaking, the multidimensional array $W(\{N_i\})$ consists of an infinite number of elements, but for computational purposes we may set an upper limit on the maximum number of i-clusters the system can possess at any time. In other words the array $W(\{N_i\})$ takes the form $W\left(0:N_1^{\max},0:N_2^{\max},\ldots,0:N_{i_{\max}}^{\max}\right)$. These values $N_1^{\max},N_2^{\max},\ldots,N_{i_{\max}}^{\max}$ should be decided by educated guess such that all of the $i_{\max}$ probability distributions in Eq. (5.7) die down to negligible levels at $N_i=N_i^{\max}$ at the end of the iterations.

A steady state is considered to have been reached when all the elements of $W(\{N_i\})$ have converged within a very small tolerance. The nucleation rates J

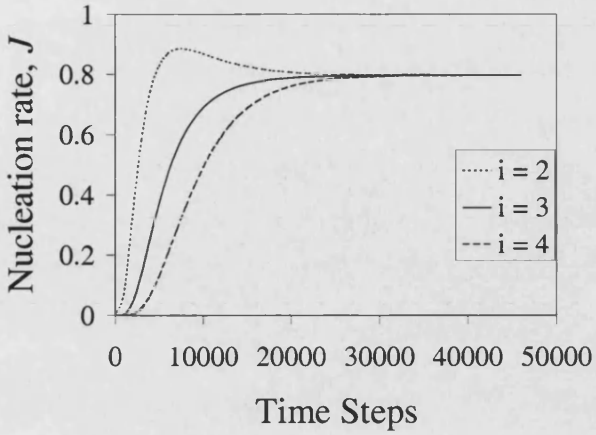

Figure 5.3: Evolution of the nucleation rate when different values of i are used in Equation (5.12). Any value of i will lead to the same nucleation rate in the steady state. Parameters used in this example were $i^* = 2$, $i_{\max} = 4$, $j_0 = \lambda = 1$ and $\xi = 1$.

with different values of i in Equation (5.12) will normally evolve differently with time, but eventually they will all converge upon a common value (Figure 5.3). This convergence of J with different values of i in fact serves as a 'double check' for ensuring that a steady state has indeed been achieved.

In Figure 5.4 we plot the classical as well as the stochastic nucleation rates obtained under different values of $i_{\max}$, with fixed values of $i^* = 2$, $j_0 = \lambda = 1$ and $\xi = 1$. As can be seen, the nucleation rate J is not very sensitive to $i_{\max}$. The stochastic J decreases slightly with increasing $i_{\max}$, but the essential message is that a value of $i_{\max} = 4$ may be trusted in order to demonstrate at least the qualitative behaviour of the system.

An example of the probability distributions $P_i(N_i)$, as defined in Equation (5.7) and calculated once the steady state has been reached, is shown in Figure 5.5. $P_1(N_1)$ is the probability distribution for the monomer population, $P_2(N_2)$ is the same for dimers, and so on. It serves to demonstrate that for a system with low

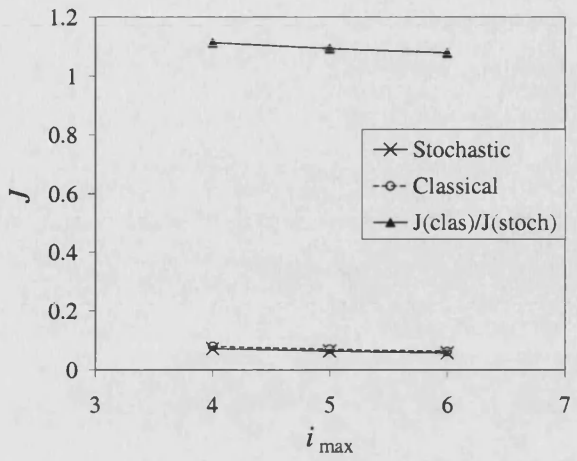

Figure 5.4: Nucleation rate as a function of $i_{\max}$ with $i^* = 2$, $j_0 = \lambda = 1$ and $\xi = 1$. The filled triangles show the ratio of classical to stochastic results for the nucleation rates, which happens to be close to unity in the example chosen here. It is reasonably safe to choose $i_{\max} = 4$, since the results obtained with a higher $i_{\max} = 6$, for instance, are approximately the same.

mean populations of the clusters, it is possible to obtain a numerical solution to the master equations by restricting ourselves to fairly small values of $N_i^{\max}$, hence making the system computationally manageable.

Figure 5.6 shows the stochastic and classical nucleation rates as a function of the particle size parameter ξ for $j_0 = \lambda = 100$. The calculation has been performed with $i^* = 2$ and $i_{\max} = 4$. Figure 5.7 shows the mean monomer population for the same system as predicted by the two models. There is a good agreement between the two models for the monomer population in this limit of $j_0 \gg 1$. The nucleation rates in Fig. 5.6 according to the two models, however, start diverging as ξ falls below 0.1. It is interesting to note that the mean monomer population between $\xi = 0.1$ and $\xi = 1$ is below unity and yet the stochastic nucleation rate does not differ considerably from its classical counterpart in this range and for these parameters.

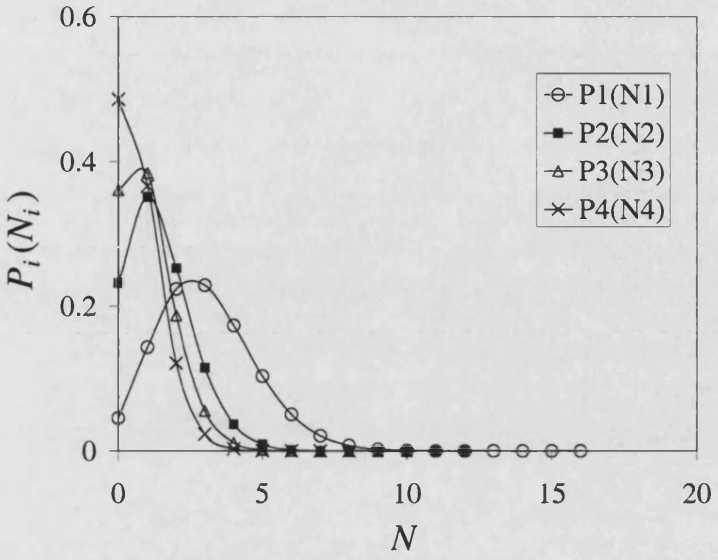

Figure 5.5: A typical example of probability distributions $P_i(N_i)$. Only values plotted at integer N are physical; the curves have been fitted as a guide to the eye. In this example, $N_1^{\max} = 16$, $N_2^{\max} = 12$ etc. were sufficient to give satisfactorily smooth probability distributions for the mean populations.

In Figure 5.8 the nucleation rate is plotted again as a function of ξ, but this time with $j_0 = \lambda = 1$, the rest of the parameters being the same as in Fig. 5.6. The mean monomer population for the same system is plotted in Figure 5.9 and now we see that the stochastic $\langle N_1 \rangle$ does differ from classical $\langle N_1 \rangle$ once ξ goes below unity. Approximately below the size $\xi = 1$, where the mean monomer population is below unity, visible difference between the classical and stochastic nucleation rates is again evident in Figure 5.8. The linear dependence of J with respect to ξ exhibited in the classical theory is lost when one deals with very small particle sizes. Note that the stochastic model gives a smaller nucleation rate, but a higher mean population of monomers than the classical prediction, since a higher nucleation rate would leave fewer monomers on the surface.

The ratios $J_{\text{classical}}/J_{\text{stochastic}}$ derived from both cases, $j_0 = \lambda = 100$ and

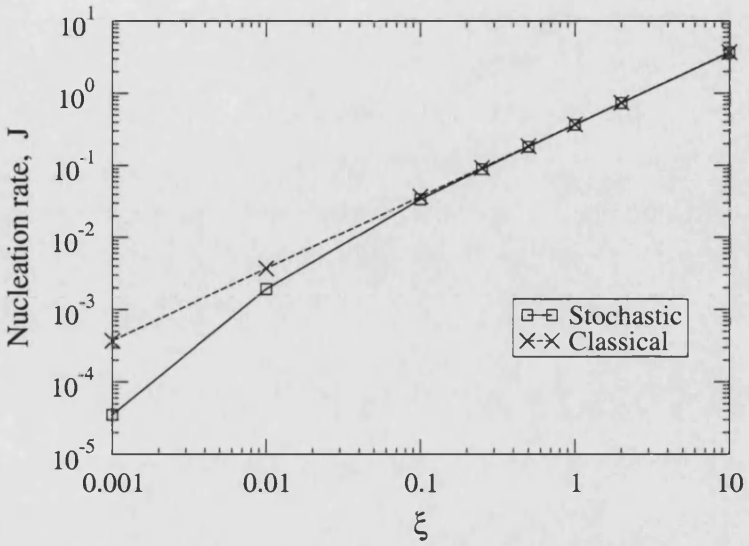

Figure 5.6: Nucleation rate as a function of the size parameter ξ for the $j_0 = \lambda = 100$ model. The prediction of rate equation approach is shown with cross signs and the squares are the results of the stochastic model presented here.

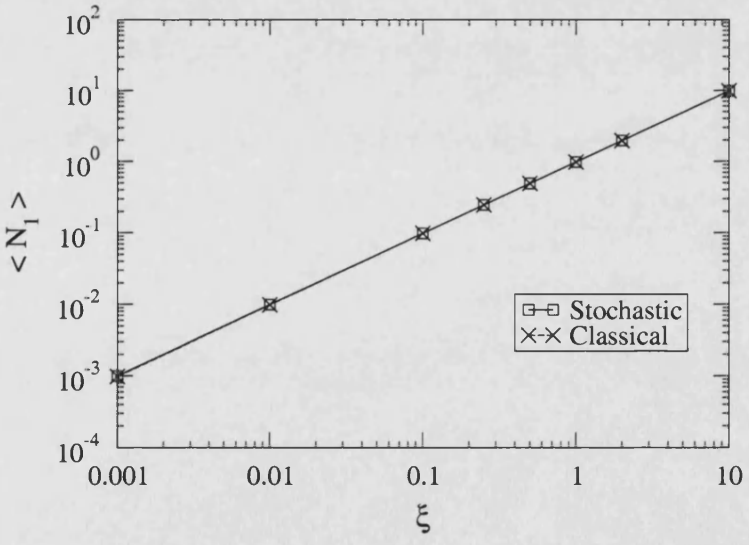

Figure 5.7: Stochastic and classical mean monomer population, $\langle N_1 \rangle$, as a function of ξ for the $j_0 = \lambda = 100$ model. Both models predict essentially the same mean populations for this choice of parameters.

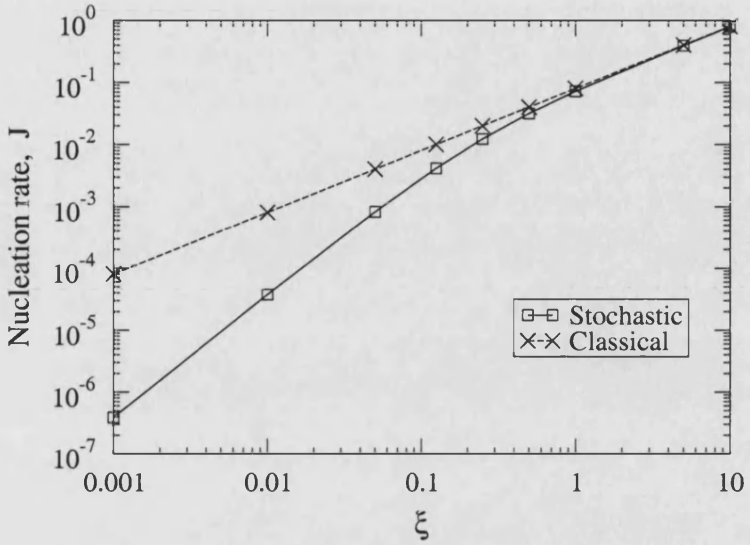

Figure 5.8: Nucleation rate as a function of ξ for the $j_0 = \lambda = 1$ model. Difference between the stochastic and classical models emerges below $\xi = 1$.

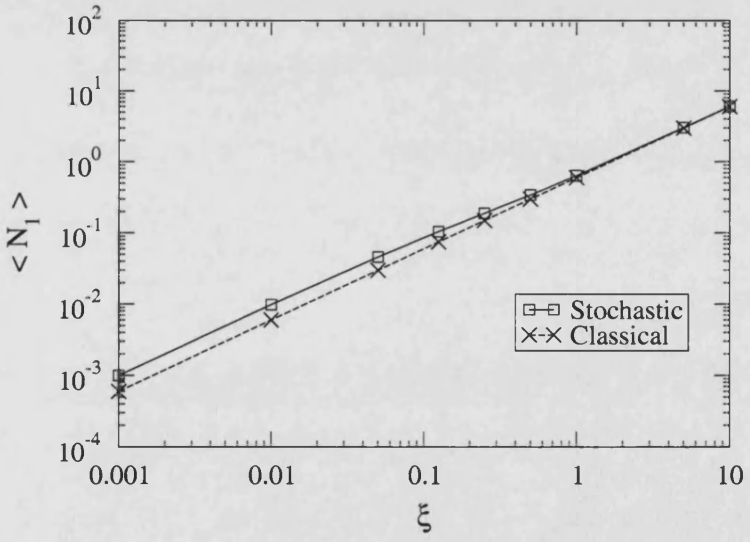

Figure 5.9: Stochastic and classical mean monomer population, $\langle N_1 \rangle$, as a function of ξ for the $j_0 = \lambda = 1$ case. Unlike the $j_0 = \lambda = 100$ case, some difference can be seen here between the mean populations according to the two models.

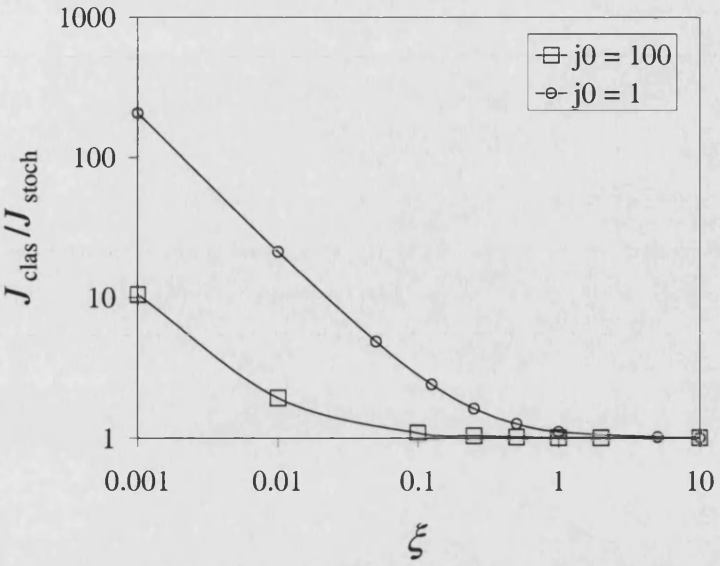

Figure 5.10: The ratio of classical versus stochastic nucleation rate calculated as a function of ξ for the $j_0 = \lambda = 100$ and $j_0 = \lambda = 1$ models.

$j_0 = \lambda = 1$, have been plotted in Figure 5.10. This is simply the factor by which the classical Becker-Döring kinetics overestimates the nucleation rate as compared with the stochastic model presented here. The overestimation grows as we look at ever smaller sizes (ξ) of the host particle. Also, the ratio is larger for the $j_0 = \lambda = 1$ calculations, compared with the $j_0 = \lambda = 100$ case. This is due partly to the fact that a large value of j_0 produces a mean monomer population closer to the classical prediction as discussed in §5.4.1.

The classical treatment requires there to be a large population of the nucleating species so as to be able to use a mean value of the populations in treating the kinetics. However, when the mean monomer population is below unity, there are instances when there are no monomers present on the surface and only by a lucky chance are there more than one monomers present. Since the classical kinetics ignores this discrete nature of the molecular species, it assumes a higher reaction rate between

the molecules, hence yielding an overestimated nucleation rate.

5.5 Conclusions

The problem of heterogeneous nucleation under conditions where the mean populations of the nucleating clusters may be of the order of unity has been studied here. The traditional rate equation approach, which treats the kinetics in terms of the mean cluster populations, is likely to fail in such limit. To investigate this, a new master equation approach has been proposed that takes into account the stochastic fluctuations in cluster populations and replaces the classical rate equations.

A method for solving the master equation numerically has been explored. The results of the model calculations performed here indicate a large difference in the nucleation rates as predicted by the stochastic and classical treatments as the nucleation site becomes very small. However, if the system size and the mean populations are large, the stochastic treatment reproduces the classical Becker-Döring kinetics.

For simplicity, only monomer attachment and detachment to the nucleating cluster has been allowed in the stochastic model here. The master equation can nevertheless be extended easily to include the loss and gain of dimers, trimers etc., solving which would clearly require a much greater deal of computational power.

Chapter 6

The Primary Nucleating Species in Dust Formation: MgO as a Candidate

6.1 Dust formation and MgO

In Chapters 3 and 5 we saw how dust particles play a crucial catalytic role in interstellar chemistry and heterogeneous mantle growth. The dust particles themselves, however, are very unlikely to form in the ultra-low density conditions of the interstellar medium (ISM). We saw how difficult it is for molecules to form, or small clusters to nucleate, in the ISM even with the help of a catalysing surface. The painstakingly long chain of reactions required to form a macroscopic dust grain from an unstable phase may be so slow in the ISM that the cluster would probably disintegrate before it shapes into a stable grain. For clusters to grow to radius of 10^{-7} m under typical interstellar conditions, the timescale required would be in terms of billions of years [71], i.e. comparable to the age of the universe itself, hence ruling out the ISM as a potential factory for cosmic dust grains.

It therefore looks more plausible that the process of dust formation takes place well before the material is thrown out in the ISM – in the denser regions of circumstellar shells around stars that have completed one full lifecycle and are losing material rapidly due to thermal pulsation, which in a sense is a sign of the star breathing its last. The relatively high temperature around stars ($\sim 10^3$ K) also makes an ideal condition for gaseous atoms to fall into the most stable molecular states before participating in dust formation. The possibility of stellar atmospheres as regions for nucleation of stardust was first pointed out in 1962 by Hoyle and Wickramasinghe [72] and today this has become a widely accepted picture.

Circumstellar dust exists around most type of stellar objects [73]. These shells are huge: if R_s is the radius of the actual star, the inner boundary of the shell normally starts at $\sim 10R_s$ and outer boundary would be at $\sim 10^4 - 10^5 R_s$ [15, p. 228]. The outer boundary is marked when the density and temperature within the shell become comparable with those of the ISM. Dust formation is believed to occur around a variety of objects such as late type giants and supergiants of spectral class M and C, novae, R CorB stars, WR-stars and S Dor variables [74]. Among these, late type giants and supergiants have a unique combination of advantages when concerned with the study of dust formation. They are quite numerous in the sky and their mass loss rate is very high; estimates suggest [75, 76, 77] that for very luminous red giants and supergiants, the mass loss rate can be as high as $10^{-4}\ M_\odot\,\mathrm{yr}^{-1}$, where $M_\odot$ denotes the mass of the Sun. This is an enormously large value when one notes that for the Sun, for example, the current mass loss rate is only about $2 \times 10^{-14}\ M_\odot\,\mathrm{yr}^{-1}$. In total, late type giants are thought to eject more than $0.3\ M_\odot\,\mathrm{yr}^{-1}$ of material into the interstellar medium in our Galaxy [73, 78, 79]. Unlike nova outbursts, giants and supergiants do not exhibit any strong

source of ultraviolet radiation, which would introduce additional complexity in the chemistry. Also, since circumstellar shells around late type giants are a significant source of infrared and microwave emission, these objects are observationally quite well studied. It is therefore sensible to study the problem of dust formation around these objects.

Carbon and oxygen turn out to be the two crucial elements that decide the chemical evolution of a cloud. Due to its particularly high bond energy (11.09 eV), the formation of carbon monoxide (CO) around late type giant stars is a one-way process. Once formed, CO is indestructible up to about 3000 K in the hostile environments around stars, unless there is any high energetic radiation to dissociate it. As the gas outflow moves away from the star, CO formation takes place between 3000 – 2000 K and it is the first molecule to materialize from the atomic phase. Since there is no UV radiation around these giants, CO formation continues unhindered until either carbon or oxygen is entirely exhausted. This fact has led astrophysicists to classify dust chemistry environments into two types: oxygen-rich and carbon-rich. If a cloud has initially more oxygen than carbon, all the carbon would have been consumed in forming CO, leaving surplus oxygen free for further chemistry. This is the case around M-stars, whose surface temperatures lie within the range of roughly 2700 – 3700 K. Similarly, for a cloud with carbon-to-oxygen ratio greater than unity, one finds an excess amount of carbon in the cloud, as around C-stars*. The large amount of carbon appears in some giant stars as a result of helium-burning and when the carbon is swept to the surface of the star, it is deposited into its circumstellar shell. As a comparison, the Solar System C/O ratio is roughly 0.5.

*There is no spectral type C among the classification of stars. C-stars are normally mixed in with G, K and M type of stars, but the name has been given to specifically indicate their carbon-richness.

Carbonaceous grain formation has been studied with quite a lot of success and the associated chemical pathway is rather well understood [74, 80, 81]. This, however, is not the case for inorganic dust formation around M-stars. Astronomers can tell us the overall chemical composition of dust in oxygen-rich clouds, but identifying the primary condensing material has proved to be less than straightforward. Interpretation of spectral features of the interstellar grains with the help of laboratory experiments [15, 82, 83, 84, 85] has brought about widespread acceptance that the main constituents of dust around M-stars are amorphous silicates. However, they cannot be the material directly nucleating from the gas phase. The formation of an amorphous dust species would have required annealing at low temperature and in stellar environments, one certainly finds high temperature. The seed nuclei therefore have to be formed in crystalline form on which further mantle growth occurs heterogeneously. It is also believed that the most abundant monomeric species that could facilitate the production of dust are Fe, Mg, SiO and H_2O. Among these, the SiO molecules has a high bond energy ($\sim$ 8.3 eV). One popular belief was that SiO crystals nucleate first and then all these four species get involved in further growth of the particle. This scenario was, however, discarded on the grounds that the major condensation phase typically occurs around late type M-stars within the temperature range of 800 – 1200 K and the nucleation temperature required for SiO would be 600 K or less [86, 87, 88, 89].

Earlier, spectroscopic analysis of the ISM suggested that some 10–30% of interstellar magnesium is locked into MgO solids [90]. Further, MgO, often known as *periclase*, is among the species that possess some of the highest dew points, which would enable it to nucleate at high temperatures. Could it then be the primary nucleating species in circumstellar clouds, some of which then transfers radially outwards

to the ISM? Köhler *et al.* investigated this possibility [91] and eventually concluded that the nucleation rate of MgO would be too small in stellar environments for it to form stable dust seeds. It appeared that MgO vapour could be supercooled significantly, leading to a lower-than-expected condensation temperature, so that it no longer stood out on the basis of its high dew point temperature. The production rate of the condensed phase was low even for large degrees of supercooling since the work of formation of the critical cluster remained high. These calculations were based on a semi-empirical potential model that took into account several factors. These included the so-called T-Rittner potential, which considers polarizability of ions due to local electric fields in small clusters, and additional terms to account for a covalent character in the Mg–O bonding.

However, it has been found that magnesium oxide requires a much more detailed potential model in order to correctly explain its experimental data ranging from the interionic distance to phonon dispersion curves. In the present work, we therefore revisit the calculations of the critical work of formation of MgO, using a potential that is a combination of the 'compressible ion model' (CIM) and the 'polarizable ion model' (PIM) that was developed by Wilson and colleagues [92, 93, 94, 95]. It is a sophisticated, transferable interionic potential designed to provide a better description of the properties of MgO clusters than any other models used in earlier studies. It not only considers the polarizability of ions, but also allows for changes in an ion's size and shape. The ions themselves are seen to be 'breathing' individually, making the model significantly more detailed than earlier models.

In the next section, the potential model of previous studies by Köhler *et al.* is briefly outlined and §6.3 contrasts this with the more accurate model of Wilson *et al.* The scheme for obtaining cluster free energy is described in §6.4, followed by

the results in §6.5. Incidentally, we explore a novel method for calculating the free energies in Appendix B.

6.2 Potential used previously

We begin by briefly noting the main features of the potential used in the earlier study of Köhler *et al.* [91]. Details such as numerical parameters are skipped here and can be found in the original publication. The main purpose here is to gain a qualitative flavour of the model used.

6.2.1 Potential for simple ionic crystals

Magnesium oxide is a highly ionic species when the interionic distance is close to equilibrium. Ionic crystals are usually described by the Born-Mayer potential [96], the simplest form of the potential energy being

$$U_{\text{B-M}} = \frac{1}{2}\sum_{\substack{\ell,j \\ \ell\neq j}} \frac{Q_\ell Q_j}{r_{\ell j}} + \frac{1}{2}\sum_{\substack{\ell,j \\ \ell\neq j}} A_{\ell j}\, exp\left(-\frac{r_{\ell j}}{\rho_{\ell j}}\right). \tag{6.1}$$

The first summation represents Coulomb interaction between charges Q_ℓ and Q_j, situated at lattice sites labelled ℓ and j respectively and whose mutual separation is $r_{\ell j}$. The second summation quantifies the repulsion due to the overlap of the electron densities of two ions situated at ℓ and j. $A_{\ell j}$ is the strength of this repulsion and $\rho_{\ell j}$ measures how steep the repulsive potential is.

6.2.2 Rittner potential

In the B–M potential, individual ions are considered to be spherical charged distributions packed together in a stable lattice configuration. For an infinitely large

lattice with high symmetry, this may hold true since electric fields in the immediate vicinity of a given ion would cancel each other out. When one deals with small particles, however, this symmetry is lost and hence each ion will experience a net electric field, which leads to the ion's polarization.

Rittner [97] allowed for this effect in alkali-halide (MX) systems of diatomics by incorporating the monopole-induced dipole interaction, the induced-dipole-induced-dipole interaction and the work required to form the induced dipole moment. The resulting potential has the form

$$U_{\text{Ritt}} = \sum_{\ell,j \text{ pairs}} \frac{Q_\ell Q_j}{r_{\ell j}} + \sum_{\ell,j \text{ pairs}} A_{\ell j} \, exp\left(-\frac{r_{\ell j}}{\rho_{\ell j}}\right) - \sum_{\ell,j \text{ pairs}} \left[\frac{Q_\ell Q_j(\alpha_\ell + \alpha_j)}{2\, r_{\ell j}^4} - \frac{2\, Q_\ell Q_j(\alpha_\ell \alpha_j)}{r_{\ell j}^7} - \frac{C_{\ell j}}{r_{\ell j}^6}\right], \quad (6.2)$$

where α_ℓ is the polarizability of the ℓ-th ion. The last term involving $C_{\ell j}$ is the London approximation to the long range van der Waals attraction, but was neglected in the work of Köhler *et al.* on the grounds that it contributes less than one percent to the total bond energy that is generated by the electrostatic forces.

It was later shown [98] that the induced polarization terms in the squared braces of Equation (6.2) can be simplified as

$$U_{\text{ind}} = -\sum_{\ell,j \text{ pairs}} \frac{Q_\ell(\boldsymbol{\mu}_j \cdot \mathbf{r}_{\ell j})}{r_{\ell j}^3}, \quad (6.3)$$

with

$$\boldsymbol{\mu}_j = \alpha_j \mathbf{E}_j = \alpha_j \sum_{\substack{k \\ k \neq j}} \left[\frac{Q_k \mathbf{r}_{kj}}{r_{kj}^3} + \frac{3\mathbf{r}_{jk}(\boldsymbol{\mu}_k \cdot \mathbf{r}_{jk})}{r_{jk}^5} - \frac{\boldsymbol{\mu}_k}{r_{jk}}\right]. \quad (6.4)$$

Here, $\boldsymbol{\mu}_j$ is the induced dipole moment at the position of ion j and is generated by the local electric field $\mathbf{E}_j$ due to all charges Q_k and induced dipoles $\boldsymbol{\mu}_k$.

6.2.3 T-Rittner potential

Further simplification of the Rittner potential came when analysis based on quantum mechanical perturbation theory suggested that consistency with this theory requires that the term proportional to $1/r_{\ell j}^7$ in Equation (6.2), together with the last two terms in the expression for $\boldsymbol{\mu}_j$ in Equation (6.4), ought to be dropped [99]. This is known as the *truncated-Rittner* (T-Rittner) potential:

$$U_{\text{T-Ritt}} = \sum_{\ell,j \text{ pairs}} \frac{Q_\ell Q_j}{r_{\ell j}} + \sum_{\ell,j \text{ pairs}} A_{\ell j}\, exp\left(-\frac{r_{\ell j}}{\rho_{\ell j}}\right) - \sum_{\ell,j \text{ pairs}} \frac{Q_\ell(\boldsymbol{\mu}_j \cdot \mathbf{r}_{\ell j})}{r_{\ell j}^3} \tag{6.5}$$

where

$$\boldsymbol{\mu}_j = \alpha_j \sum_{\substack{k \\ k \neq j}} \frac{Q_k\, \mathbf{r}_{kj}}{r_{kj}^3}. \tag{6.6}$$

6.2.4 Modified T-Rittner potential

The potential described so far (Eq. (6.5)) was developed primarily to describe the bonding of group I-VII comounds. Group II-VI compounds such as MgO have similar molecular characteristics, but they are not as purely ionic as the I-VII materials. The relatively moderate difference in the electronegativity of ions in MgO suggests a significant contribution of covalent bonding [100]. The Morse potential for covalently bound diatomic molecules is

$$U_{\text{Morse}} = \mathcal{A} \exp\left(-2\frac{r_{\ell j}}{\rho_{\ell j}}\right) - \mathcal{B} \exp\left(-\frac{r_{\ell j}}{\rho_{\ell j}}\right) \tag{6.7}$$

with

$$\mathcal{A} = D_0 \mathrm{e}^{2a}, \quad \mathcal{B} = 2D_0 \mathrm{e}^{a} \quad \text{and} \quad \rho = r_0/a. \tag{6.8}$$

Here D_0 is the depth of the potential, i.e. the dissociation energy reduced by the contribution of the zero-point energy of the vibrations. r_0 denotes the equilibrium

distance of ions and a is a free parameter that is usually fitted by requiring

$$\mu\omega^2 = \left.\frac{\mathrm{d}^2U(r)}{\mathrm{d}r^2}\right|_{r=r_0} = 2\,D_0\left(\frac{a}{r_0}\right)^2, \tag{6.9}$$

where μ is the reduced mass and ω is the vibrational frequency of the molecule.

The second term in the Morse potential describes covalent attraction. The quantity $\mathcal{B}$ vanishes in a large ionic solid, but is nonzero for small clusters. It is therefore dependent on the number of MgO monomers in the cluster. Equation (6.7) also has a repulsive term similar to the second term in Born-Mayer potential in Equation (6.1). Therefore, in their final potential, Köhler *et al.* replace the $A_{\ell j}$ term with the $\mathcal{A}$ term. Furthermore, rather than using bare charge Q_ℓ, an effective charge Q_ℓ^{eff} was used in their calculations due to the lower electronegativity difference of the ions [101]. Hence, merging Equations (6.1), (6.5) and (6.7), we arrive at the full potential that was employed in the previous work:

$$U_{\mathrm{Koh}}(N) = \sum_{\ell,j\text{ pairs}} \mathcal{A}\mathrm{e}^{-\frac{2r_{\ell j}}{\rho}} - \sum_{\ell,j\text{ pairs}} \mathcal{B}\mathrm{e}^{-\frac{r_{\ell j}}{\rho}} + \sum_{\ell,j\text{ pairs}} \frac{Q_\ell^{\mathrm{eff}}Q_j^{\mathrm{eff}}}{r_{\ell j}} - \sum_{\ell,j\text{ pairs}} \frac{Q_\ell^{\mathrm{eff}}(\boldsymbol{\mu}_j\cdot\mathbf{r}_{\ell j})}{r_{\ell j}^3}, \tag{6.10}$$

where $\boldsymbol{\mu}_j$ is given by Equation (6.6), but with Q_k^{eff} replacing Q_k.

6.3 Potential used in the present work

We now go through the key features of the potential model that was used for the present calculations. It is a complex model that was developed through a series of publications [92, 93, 94, 95] in order to include a number of effects. The most radical of these is that individual ions are treated as compressible balls of charge distribution. This is a significant departure from the models described in the previous section, in which electron clouds of two neighbouring ions were allowed to overlap each other, but the ions maintained their original identity in terms of ionic

radii. The only way ionic shape deformation was allowed in the Rittner potential, for example, was through polarization effects in finite size (small) clusters. It is therefore often termed a 'rigid ion model' (RIM). In the compressible ion model (CIM) [92], the size of an ion is allowed to change via compression of its charge distribution, effected by the neighbouring ions. Polarization effects are then added through the polarizable ion model (PIM) [93] and to further the refinement, there is also a scheme for allowing aspherical shape deformations of the ion [95].

This model is also more flexible than the Shell Model [102], which was traditionally used to describe effects due to an ion's environment. By chance, ionic compression features do appear, in a sense, in "breathing-shell" models [103, 104]. However, these models still require a large number of parameters to be fixed by fitting them to experimental data obtained for some specific material and for a given phase. It is then very difficult to ascertain that the same parameters would hold true for a different phase of the same species. In contrast, the CIM + PIM representation relies on parameterisation with the aid of *ab initio* electronic structure calculations, making it fully transferable between chemically related species as well as between different phases.

6.3.1 Compressible ion model

As the starting point of the CIM, the Born-Mayer potential is again used, but with a more elaborate form:

$$
\begin{aligned}
U_{\text{B-M}} &= \frac{1}{2}\sum_{\substack{\ell,j \\ \ell\neq j}} \frac{Q_\ell Q_j}{r_{\ell j}} + \frac{1}{2}\sum_{\substack{\ell,j \\ \ell\neq j}} A_{\ell j}\, exp\left[-a_{\ell j}(r_{\ell j} - \sigma_\ell - \sigma_j)\right] \\
&\quad - \frac{1}{2}\sum_{\substack{\ell,j \\ \ell\neq j}} \left[\frac{C_6^{\ell j}}{r_{\ell j}^6} f_6^{\ell j}(r_{\ell j}) + \frac{C_8^{\ell j}}{r_{\ell j}^8} f_8^{\ell j}(r_{\ell j})\right] \qquad (6.11) \\
&= U_{\text{Cou}} + U_{\text{rep}} + U_{\text{disp}},
\end{aligned}
$$

i.e. the potential is a sum of Coulombic attraction, short range repulsion and an additional dispersion term. The repulsive potential in the second summation now includes ionic radii σ_ℓ and σ_j compared to Equation (6.1) and the range parameter, $a_{\ell j}$, characterises the shape of the charge density. $C_6^{\ell j}$ and $C_8^{\ell j}$ are the dipole-dipole and dipole-quadrupole dispersion parameters respectively, relevant to interaction between species ℓ and j. These can be derived from *ab initio* calculations or by experiments [105]. $f_n^{\ell j}(r_{\ell j})$ are the damping functions to characterise the effect of the overlap of electron densities of ions ℓ and j and are represented by the so-called Tang-Toennies functions [106]. When the overlap between two ions' wavefunctions is negligible, the damping functions are unity, but they drop towards zero as the overlap becomes significant. Hence the $f_n^{\ell j}(r_{\ell j})$ reduce the effect of the dispersion terms in overlapping electron shells.

It is the repulsive part of the potential, U_{rep}, that requires attention in order to allow for compression effects. In the CIM, the repulsive potential is a function of not only the ionic positions ($\{\mathbf{r}_\ell\}_{\ell=1,N}$), where N is the total number of ions in the system, but also the change in the instantaneous ionic radii, $\delta\sigma_\ell$, with respect to some reference values $\bar{\sigma}_\ell$. Further, U_{rep} can be seen to be arising from two contributions: (i) the overlap between the electron clouds of two neighbouring ions (U_{ov}) and (ii) the energy it costs to deform each ion's electron density by the amount $\delta\sigma_\ell$, denoted

by U_{self}. In other words,

$$U_{\text{rep}}(\{\mathbf{r}_\ell, \delta\sigma_\ell\}_{\ell=1,N}) = U_{\text{ov}}(\{\mathbf{r}_\ell, \delta\sigma_\ell\}_{\ell=1,N}) + U_{\text{self}}(\{\mathbf{r}_\ell, \delta\sigma_\ell\}_{\ell=1,N}). \tag{6.12}$$

The simplest form for U_{ov} is

$$U_{\text{ov}}(\{\mathbf{r}_\ell, \delta\sigma_\ell\}_{\ell=1,N}) = \sum_{\ell,j \text{ pairs}} A_{\ell j} \exp\Big(- a_{\ell j} \left[r_{\ell j} - (\bar{\sigma}_\ell + \delta\sigma_\ell) - (\bar{\sigma}_j + \delta\sigma_j) \right] \Big). \tag{6.13}$$

The Mg^{2+} ion is assumed to be electronically rigid compared to O^{2-} owing essentially to the electronegativity of the latter and also because the polarizability of Mg^{2+} is some 3 orders of magnitude smaller than that of O^{2-} in the crystal. A suitable form for U_{self} for an oxide ion is

$$U_{\text{self}}(\{\delta\sigma_\ell\}_{\ell=1,N}) = \sum_\ell D \left(e^{\beta\, \delta\sigma_\ell} + e^{-\beta\, \delta\sigma_\ell} \right), \tag{6.14}$$

where $2D$ is the energy of the reference ion without any compression ($\delta\sigma_\ell = 0$) and can be taken as the second electron affinity of oxygen [107]. The parameter β is such that $D\beta^2$ is the harmonic force constant that resists the ion's breathing originating from the compression-decompression process.

6.3.2 Polarizable ion model

In this model, polarization effects are represented in a more realistic way than through the T-Rittner method. One has to consider not only polarization due to electric fields and field gradients due to the charges and dipole moments of neighbouring ions, but also the fact that as an ion's immediate neighbour moves off its lattice site, the shape of the confining potential changes. This in turn leads to additional short-range effects in the polarization potential energy. Furthermore, it turns out that the "covalent" bond contribution suspected in the earlier work (see §6.2.4) has been shown to be a manifestation of thoroughly defined polarization effects [108, 109].

In the PIM [95, 110], additional degrees of freedom are assigned to each ion ℓ alongside its position $\mathbf{r}_\ell$ and the instantaneous change in its radius discussed in the CIM. These are the electric dipole moment $\boldsymbol{\mu}_\ell = \{\mu_\ell^\alpha\}_{(\alpha=x,y,z)}$ and the quadrupole moment $\boldsymbol{\theta}_\ell = \{\theta_\ell^{\alpha\beta}\}_{(\alpha\beta=xx,yy,zz,xy,xz,yz)}$. In other words, $\boldsymbol{\mu}_\ell$ has three components, whereas $\boldsymbol{\theta}_\ell$ has six, although $\theta_\ell^{\alpha\beta} = \theta_\ell^{\beta\alpha}$ implies that only five components of $\boldsymbol{\theta}_\ell$ are independent. The Greek symbols α, β etc., when appearing as superscripts in this discussion, will take on any of the coordinate values x, y or z. The dipole and quadrupole moments are now related to the local electric field via the relations (cf. Equation (6.4))

$$\mu_\ell^\alpha = \alpha^{\alpha\beta} E^\beta(\mathbf{r}_\ell) + \frac{1}{3} B^{\alpha\beta,\gamma\delta} E^\beta(\mathbf{r}_\ell) E^{\gamma\delta}(\mathbf{r}_\ell) \tag{6.15}$$

and

$$\theta_\ell^{\alpha\beta} = \frac{1}{2} B^{\alpha\beta,\gamma\delta} E^\gamma(\mathbf{r}_\ell) E^\delta(\mathbf{r}_\ell) + C^{\alpha\beta\gamma\delta} E^{\gamma\delta}(\mathbf{r}_\ell). \tag{6.16}$$

Here E^α and $E^{\alpha\beta}$ are components of the electric field and field gradient respectively. $\boldsymbol{\alpha}$ and $\mathbf{C}$ are the dipole and quadrupole polarizabilities and $\mathbf{B}$ is the dipole-dipole-quadrupole hyperpolarizability. After some lengthy analysis, it is possible to show [111] that for a spherical ion, the components of $\boldsymbol{\alpha}$, $\mathbf{C}$ and $\mathbf{B}$ are determined by a single number.

The route to a generalised polarization potential is quite involved, but essentially the polarization contribution to the potential energy of a system of ions is given by

$$U_{\text{pol}} = \sum_{\ell,j \text{ pairs}} \Big[f^{(1)}(r_{\ell j})\, T^{\alpha}_{\ell j}(Q_j \mu^{\alpha}_{\ell} - Q_{\ell}\mu^{\alpha}_{j}) - T^{\alpha\beta}_{\ell j}\mu^{\alpha}_{\ell}\mu^{\beta}_{j} + \frac{1}{3} T^{\alpha\beta}_{\ell j} f^{(2)}(r_{\ell j}) \left(Q_{\ell}\theta^{\alpha\beta}_{j} + Q_j \theta^{\alpha\beta}_{\ell}\right) + \frac{1}{3} T^{\alpha\beta\gamma}_{\ell j}(\mu^{\alpha}_{\ell}\theta^{\beta\gamma}_{j} - \mu^{\alpha}_{j}\theta^{\beta\gamma}_{\ell}) + \frac{1}{9} T^{\alpha\beta\gamma\delta}_{\ell j}(\theta^{\alpha\beta}_{\ell}\theta^{\gamma\beta}_{j}) \Big] + k_1 \sum_{\ell}^{N}(\mu_{\ell})^2 + k_2 \sum_{\ell}^{N} \mu^{\alpha}_{\ell}\theta^{\alpha\beta}_{\ell}\mu^{\beta}_{\ell} + k_3 \sum_{\ell}^{N} \theta^{\alpha\beta}_{\ell}\theta^{\alpha\beta}_{\ell} + k_4 \sum_{\ell}^{N}(\mu_{\ell})^4. \quad (6.17)$$

The '**T** tensors' appearing here can be expressed as [111]

$$T^{\alpha\beta\gamma\delta\ldots}_{\ell j} = \nabla^{\alpha}\nabla^{\beta}\nabla^{\gamma}\nabla^{\delta}\ldots \frac{1}{r_{\ell j}}. \quad (6.18)$$

These tensors are symmetric in all suffixes and

$$\nabla^{\alpha}\nabla^{\alpha}\, r^{-1}_{\ell j} = (\nabla)^2\, r^{-1}_{\ell j} = 0, \quad (6.19)$$

i.e. a repeated Greek suffix reduces any **T** to zero.

The first five terms in the squared braces of Equation (6.17) represent charge-dipole, dipole-dipole, charge-quadrupole, dipole-quadrupole and quadrupole-quadrupole interactions respectively. The radial functions $f^{(1)}(r_{\ell j})$ and $f^{(2)}(r_{\ell j})$ modify the interaction of charges with induced dipoles and quadrupoles respectively. These functions have a short-range effect only and are effective over length scales of nearest-neighbour separations. In practice, they are chosen to be of Tang-Toennies dispersion damping function form,

$$f^{(n_k)}(r_{\ell j}) = 1 - \mathrm{e}^{-br_{\ell j}} \sum_{k=0}^{n_k} \frac{(br_{\ell j})^k}{k!}, \quad (6.20)$$

where b is the so-called 'short-range damping parameter', which is the reciprocal of the length scale over which the damping is effective. This length scale would be comparable to the $\bar{\sigma}$ introduced in the CIM. The terms with k_j $(j = 1, 2, 3, 4)$ in Equation (6.17) represent the energy required to polarize the ion. The parameters k_j

are harmonic and anharmonic force constants, which are related to polarizabilities α, C and B as described in Ref. [110].

It has been noted [95] that for MgO, the quadrupolar effects are negligibly small in explaining phonon dispersion curves. Hence, if one considers only the case of dipole polarization, Equation (6.17) can be written in a more compact form,

$$\begin{aligned} U_{\text{pol}}(\{\mathbf{r}_\ell, \boldsymbol{\mu}_\ell\}_{\ell=1,N}) &= \sum_{\ell,j \text{ pairs}} f^{(1)}(r_{\ell j})\, Q_\ell \mathbf{T}^{(1)}(\mathbf{r}_{\ell j}) \cdot \boldsymbol{\mu}_j \\ &+ \sum_{\ell,j \text{ pairs}} \boldsymbol{\mu}_\ell \mathbf{T}^{(2)}(\mathbf{r}_{\ell j}) \cdot \boldsymbol{\mu}_j \\ &+ \sum_{\ell} (1/2) k_\ell \, \mu_\ell^2, \end{aligned} \tag{6.21}$$

where $\mathbf{T}^{(1)}$ and $\mathbf{T}^{(2)}$ are the charge-dipole and dipole-dipole interaction tensors.

6.3.3 Aspherical ion model

The expressions for the repulsive potential given in Equations (6.13) and (6.14) apply well for spherical compression of the ion. However, they still do not accurately reproduce desired phonon dispersion curves for MgO. For a better description of the interactions, one needs to consider aspherical ion deformation. For instance, an oxide ion is inhomogeneously compressed by a number of magnesium ions from different directions (maximum six neighbours) and the resulting shape of the O^{2-} ion would not be spherical. To incorporate aspherical compression, some modification of the repulsive potential is required. This is done by including yet more degrees of freedom (DOF) in terms of the parameters of dipole and quadrupole symmetry, $\boldsymbol{\nu}_\ell$ (3 DOF), and $\boldsymbol{\kappa}_\ell$ (5 DOF), which control the shape of the ion's repulsive wall. The 'aspherical ion model' (AIM) is an extension of the CIM in which U_{ov} and U_{self} take more complex forms. The cation-anion part of U_{ov} is generalised to

$$U_{\rm ov}(\{\mathbf{r}_\ell, \delta\sigma_\ell, \boldsymbol{\nu}_\ell, \boldsymbol{\kappa}_\ell\}_{\ell=1,N})$$
$$= \sum_{\ell\in\text{anion}} \sum_{j\in\text{cation}} A_{-+}\exp\Big(-a_{-+}\Big[r_{\ell j} - (\bar{\sigma}_\ell + \delta\sigma_\ell) - (\bar{\sigma}_j) - \mathbf{S}^{(1)}(\mathbf{r}_{\ell j})\cdot\boldsymbol{\nu}_\ell - \mathbf{S}^{(2)}(\mathbf{r}_{\ell j})\cdot\boldsymbol{\kappa}_\ell\Big]\Big), \quad (6.22)$$

where $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ are interaction tensors whose elements are given by

$$S^{(1)}_\alpha(\mathbf{r}) = r^\alpha/r \qquad \text{and} \qquad S^{(2)}_{\alpha\beta}(\mathbf{r}) = 3\,r^\alpha r^\beta/(r)^2 - \delta^{\alpha\beta}. \quad (6.23)$$

Here $\delta^{\alpha\beta}$ is the Kronecker delta. The self energy of the CIM is generalised such that the energy required to cause the shape deformations of dipolar and quadrupolar symmetry is also considered,

$$U_{\rm self}(\{\delta\sigma_\ell, \boldsymbol{\nu}_\ell, \boldsymbol{\kappa}_\ell\}_{\ell=1,N_{\rm anion}}) = \sum_{\ell\in\text{anion}} \Big[D\,(e^{\beta\,\delta\sigma_\ell} + e^{-\beta\,\delta\sigma_\ell}) + (e^{\xi^2|\boldsymbol{\nu}^\ell|^2} - 1) + (e^{\eta^2|\boldsymbol{\kappa}^\ell|^2} - 1)\Big], \quad (6.24)$$

where ξ and η are two "force constants" that can be chosen by noting high symmetry points on the phonon dispersion curve. Here we have assumed the cations Mg^{2+} to be non-compressible and the asphericity effect on the O^{2-}–O^{2-} interactions has also been neglected since these effects have been found not to be influential in reproducing the observed phonon curves [95]. The variables $\boldsymbol{\nu}_\ell$ and $\boldsymbol{\kappa}_\ell$ minimise the total repulsive energy at each step of the molecular dynamics simulation.

6.3.4 Full potential for MgO

To recap, the CIM makes a number of modifications to the repulsive part of the potential in the Born-Mayer expression (6.12). In its simple form, the second term of (6.12) is replaced by Equations (6.13) + (6.14), but if one considers details of dipolar and quadrupolar shape distortions, a combination of (6.22) + (6.24) would be more accurate. In addition, one needs to include the polarization potential, whose

full form is given by Equation (6.17), although in the case of MgO the simplified version (6.21) may well serve the purpose.

Hence the full potential for MgO, including all the compression, polarization and aspherical deformation effects, now reads

$$U_{\text{CIM+PIM}} = \sum_{\ell,j \text{ pairs}} \frac{Q_\ell Q_j}{r_{\ell j}} - \sum_{\ell,j \text{ pairs}} \left[\frac{f_6^{\ell j}(r_{\ell j}) C_6^{\ell j}}{r_{\ell j}^6} + \frac{f_8^{\ell j}(r_{\ell j}) C_8^{\ell j}}{r_{\ell j}^8} \right] + U_{\text{rep}} + U_{\text{pol}}, \quad (6.25)$$

where

$$\begin{aligned} U_{\text{rep}} = \sum_{\ell \in \text{anion}} \sum_{j \in \text{cation}} A_{-+} \exp\Big(- a_{-+} \Big[r_{\ell j} - (\bar{\sigma}_\ell + \delta\sigma_\ell) - (\bar{\sigma}_j) \\ - \mathbf{S}^{(1)}(\mathbf{r}_{\ell j}) \cdot \boldsymbol{\nu}_\ell - \mathbf{S}^{(2)}(\mathbf{r}_{\ell j}) \cdot \boldsymbol{\kappa}_\ell \Big] \Big), \\ + \sum_{\ell \in \text{anion}} \Big[D \left(e^{\beta\, \delta\sigma_\ell} + e^{-\beta\, \delta\sigma_\ell} \right) + \left(e^{\xi^2 |\boldsymbol{\nu}^\ell|^2} - 1 \right) + \left(e^{\eta^2 |\boldsymbol{\kappa}^\ell|^2} - 1 \right) \Big]. \quad (6.26) \end{aligned}$$

and U_{pol} is given by Equation (6.17).

6.4 Free energy calculation

Equipped with this potential model, it is now possible to perform free energy calculations on MgO clusters through the equations outlined in Chapter 4. An inspection of the relevant equations, particularly Equations (4.29), (4.30) and (4.36), shows that in addition to the vapour density of MgO monomers, one needs to know the system potential energy at the mean atomic positions, $U_0(i)$, the mean ionic separations to calculate the moments of inertia, I_1, I_2 and I_3, plus the vibrational frequencies, ω_k, within the clusters. To obtain these quantities, the CIM + PIM potential was implemented using molecular dynamics (MD) simulation developed by Wilson *et al.*, with the parameters appearing in the potential model chosen as in Refs. [92, 93, 94, 95].

In the bulk solid phase, MgO displays simple cubic lattice structure of B1 type similar to NaCl (rocksalt), although it has been shown that for clusters smaller than

32 MgO monomers, nanotubes of stacked hexagons are energetically more favourable [94]. To start with, in the MD simulation, some nominal lattice structure is chosen for a given cluster size and the structure is then relaxed by performing energy minimisation. In such a configuration, U_0 is the total energy of the cluster since the crystal has no kinetic energy and the interionic distances can also be measured easily because of the lack of vibrational motion.

To obtain the vibrational frequencies, we make use of the dynamical matrix of force constants. This was calculated in the molecular dynamics by displacing each ion of the relaxed lattice structure by a small amount one by one and measuring the force it experiences as a function of the displacement with respect to the relaxed position. It is assumed that the mutual force $\mathfrak{F}_{\ell j}$ experienced between the j-th and the ℓ-th ion due to the displacement of the j-th ion by an amount δx_j is governed by Hooke's law,

$$\mathfrak{F}^{\alpha}_{\ell j} = -k^{\alpha\beta}_{\ell j}\delta x^{\beta}_{j}, \tag{6.27}$$

where $k^{\alpha\beta}_{\ell j}$ is the spring constant. α and β here go from 1 to 3 and label the components of the force and displacement. In a system of N ions, each having three degrees of freedom, there will be $3N$ spring constants. Hence the order of the dynamical matrix, which is composed of these spring constants, will be $3N$. Equivalently, in the case of a crystal with i molecules of MgO, the order will be $6i$. One then finds the eigenvalues λ_k of the mass-weighted dynamical matrix

$$\mathbf{D}' = \mathbf{M}^{-1/2}\,\mathbf{D}\,\mathbf{M}^{-1/2}, \tag{6.28}$$

where $\mathbf{D}$ is the dynamical matrix and $\mathbf{M}$ is a diagonal matrix containing the masses of the $6i$ oscillators on its diagonal [112].

The dynamical matrix, and hence $\mathbf{D}'$, provides $6i$ eigenvalues. However, six of them will, in principle, be zero because the modes corresponding to the entire

crystal's translation in three dimensions plus the rotation with respect to the three principal axes cannot contribute to the vibrational modes. Hence we are left with $6i - 6$ non-zero eigenvalues of $\mathbf{D}'$, except in the case of an isolated MgO monomer, which has one non-zero eigenvalue. The vibrational frequencies ω_k are related to these eigenvalues by

$$\omega_k = \sqrt{\lambda_k}, \tag{6.29}$$

which can be used to obtain the vibrational free energy according to Equation (4.26), so that the total free energy of a cluster can be calculated using Equation (4.29).

6.5 Results and discussion

As mentioned earlier, condensation of gaseous species around these stars is believed to occur in the temperature range of 800 – 1200 K and in the majority of cases, it is above 1000 K. We now have all the quantities required for calculating the free energies. A number of possible cluster geometries containing up to 24 MgO monomers were considered, the relaxed structures of which are shown in Figure 6.1. By calculating the free energy of each cluster as well as that of the monomer as discussed in Chapter 4, the effective work of formation of these clusters (Equation (4.36)) and the cluster number densities (Equation (4.32)) can be evaluated as a function of their size i.

The value of initial MgO vapour concentration was chosen to be $\rho_1 = 10^6$ m^{-3}, which is a typical value in shells surrounding M-stars [85, 91]. In a circumstellar shell outflow, this would not necessarily be the number density of MgO free monomers, for a vapour moving away from the central star will cool gradually and molecular clusters will be formed, consuming the free monomers and reducing their number

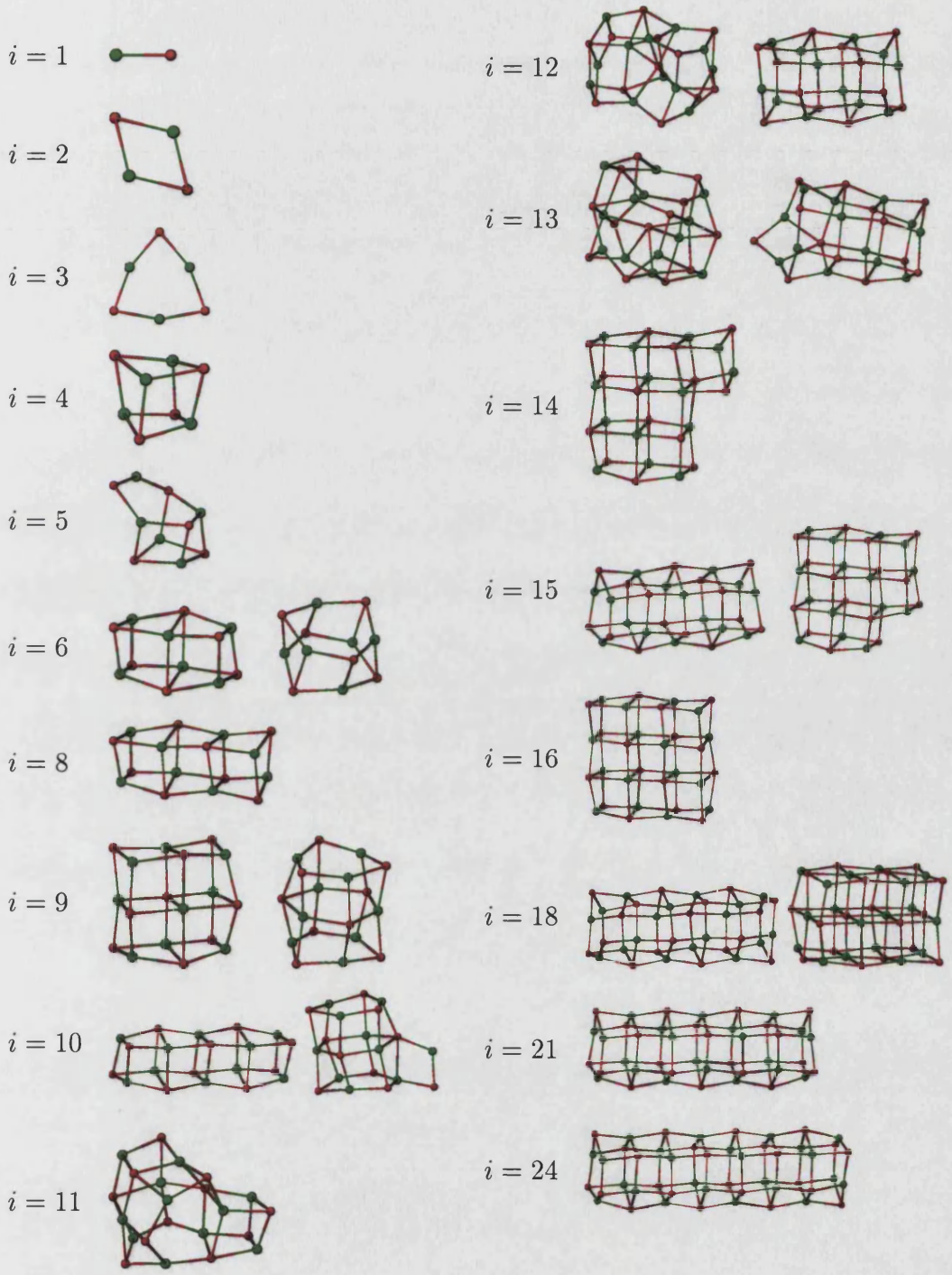

Figure 6.1: Various cluster geometries studied. Green balls are Mg^{2+} ions and red balls are O^{2-}.

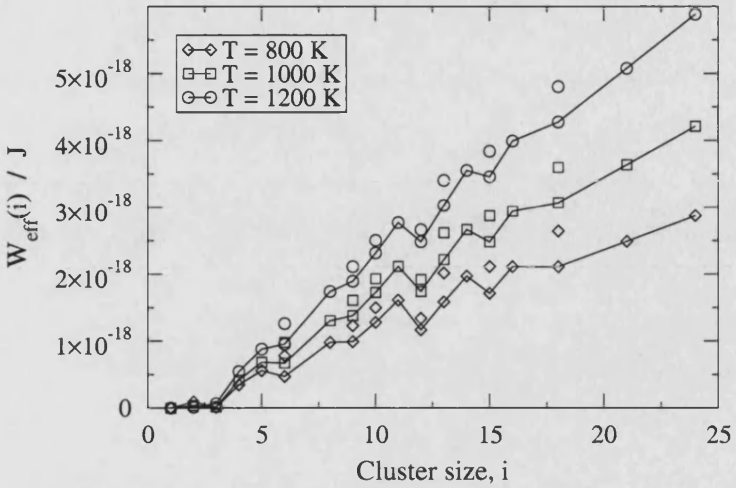

Figure 6.2: Effective work of formation, $\mathcal{W}_{\text{eff}}$, as a function of the MgO cluster size for various temperatures and a monomer concentration of 10^6 m^{-3}. The full line is drawn considering that the cluster which requires the least work of formation will be the most favoured among each size.

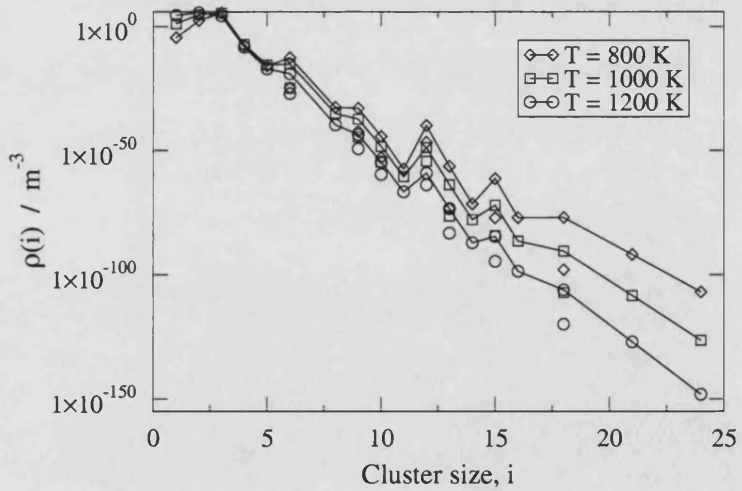

Figure 6.3: Distribution of the number densities of i-clusters for various temperatures and a monomer concentration of 10^6 m^{-3}. The full line connects the most abundant clusters of each size.

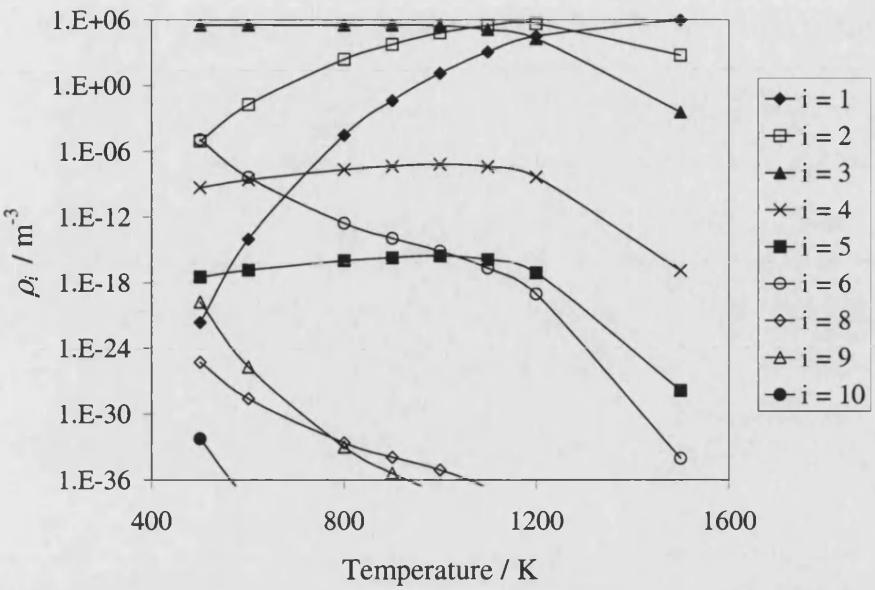

Figure 6.4: Number densities of various MgO cluster sizes as a function of temperature, with a fixed monomer concentration of 10^6 m^{-3}.

density. Hence, an iterative method was used which ensured that the total density of MgO units remained constant even after considering the number densities of all the cluster sizes considered. We do this by taking a trial value of monomer density, ρ_v, then evaluate the distribution ρ_i and calculate $\rho_{\text{tot}} = \rho_v + 2\rho_2 + 3\rho_3 + \ldots + i_{\max}\rho_{i_{\max}}$ using Equation (4.32). If ρ_{tot} does not match the desired value of ρ_1, we choose an improved trial value of ρ_v and continue to perform the calculation in this way until $\rho_{\text{tot}} \approx \rho_1$ within a small tolerance.

The effective work of formation obtained in this way for three different temperatures covering the range of interest is shown in Figure 6.2 and Figure 6.3 shows the corresponding distribution ρ_i. No distinct peak is visible in $\mathcal{W}_{\text{eff}}(i)$ and its value quickly rises to hundreds of k_BT, which indicates a lack of nucleation of MgO clusters in the given temperature range. This is also evident from the extremely small number densities for most cluster sizes, except for $i = 2$ and $i = 3$. The peak around

these two sizes in Figure 6.3 is a little unusual and suggests that MgO is an associative species, which prefers remaining in clusters of $(MgO)_2$ and $(MgO)_3$ rather than as free monomers, at least under the circumstellar conditions mentioned above.

Figure 6.4 shows how the number densities of clusters evolve as the circumstellar outflow moves away from the central star and cools down, assuming a constant vapour density. At around 1500 K, only monomers are predominant, but as the temperature drops, their concentration drops as $(MgO)_2$ and $(MgO)_3$ start to dominate. The total number density of MgO units remains fixed at 10^6 m^{-3} during the whole process.

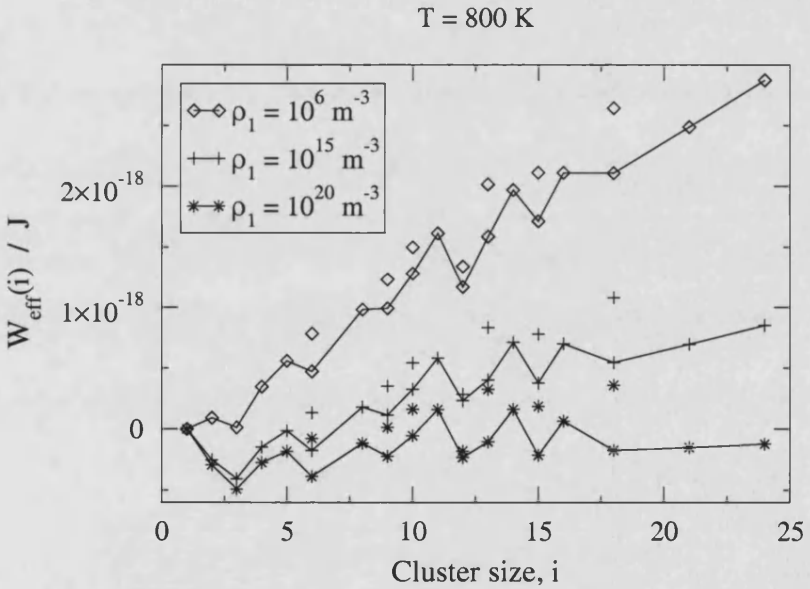

Figure 6.5: Effective work of formation, $\mathcal{W}_{\text{eff}}$, as a function of the MgO cluster size for various monomer concentrations, all at a fixed temperature $T = 800$ K. The full line joins the clusters that require the least work of formation for each size.

What if the initial number density of MgO monomers were to be significantly greater than 10^6 m^{-3}? Figure 6.5 shows the work of formation curves for two arbitrarily chosen higher values, $\rho_1 = 10^{15}$ m^{-3} and $\rho_1 = 10^{20}$ m^{-3}, each at the tem-

perature $T = 800$ K, which is the lower limit of the temperature range of interest. For comparison, the $T = 800$ K curve from Figure 6.2 has also been reproduced with these. The curve for $\rho_1 = 10^{15}$ m^{-3} does seem to be heading towards a peak value at some cluster size soon after $i = 25$, but the work of formation has already reached several dozens of k_BT in the curve shown. The nucleation rate per unit volume, J_V, is proportional to $\exp(-\mathcal{W}^*/k_BT)$, where $\mathcal{W}^*$ is the peak value of $\mathcal{W}(i)$ (Equation (4.31)). Hence nucleation in such a system would be expected to be an extremely inefficient process. Comparing this with the curve for $\rho_1 = 10^{20}$ m^{-3}, for which $\mathcal{W}^*$ appears to be very small (perhaps negative), it can be said that notable nucleation would only occur for initial vapour densities well above 10^{15} m^{-3}. Such MgO monomer concentrations are, however, atypical of circumstellar shells around oxygen-rich stars. Hence MgO ought to be discarded as a candidate for the primary nucleating in stellar winds in those environments.

6.6 Effect of associative species on nucleation

It must be remembered here that some of the equations regarding the nucleation process in Chapter 4 were built on the assumption that only monomers (single MgO molecules in this case) are allowed to attach and detach to and from a given cluster. If $(MgO)_2$ and $(MgO)_3$ concentrations overwhelm the MgO monomer abundance in the circumstellar shell, some of those equations would require modification in this light. The expressions for the nucleation rate J (and hence J_V) as written in Chapter 4 apply only when monomers outnumber any other cluster size concentrations during the entire nucleation process. For example, Equation (4.9) depends on the rate at which monomers attach themselves to the cluster of the critical size i^*. Intrinsic in this is the assumption that the concentration of monomers in the cloud is so much

greater than that of other clusters that the attachment/detachment of the latter will be too rare and would make little difference to the nucleation rate. However, in the case of circumstellar MgO, this assumption will need to be replaced by the fact that $(MgO)_2$ and $(MgO)_3$ will attach far more frequently than the monomers and the relevant equations for the nucleation rate ought to be modified accordingly.

Nevertheless, the equations for the work of formation of an i-cluster, $\mathcal{W}(i)$, will remain unaffected by this, since in that case we are merely calculating the free energy of an i-cluster with respect to that of a monomer. It simply indicates the 'effort' required to reach the size i from a monomer. Even if $(MgO)_2$ and $(MgO)_3$ were in overwhelming proportions, they would have still required to go through this effort and only then can the higher size clusters follow. $\mathcal{W}(i)$ therefore for any value of i does not depend on whether there are more monomers than other cluster concentrations. The population densities ρ_i are obtained from the work of formation, so they will also be unaffected by the associative nature of MgO. Hence, the results presented in this Chapter are indeed valid. Of course, no peak in the curve for $\mathcal{W}_{\text{eff}}(i)$ was observed in the calculations for a realistic value of vapour density, so the question of evaluating the nucleation rate did not arise.

6.7 Conclusion

In this Chapter we have studied the homogeneous nucleation of MgO clusters and explored the possibility that it could be the primary nucleating material in circumstellar shells around oxygen-rich M-stars. To do this, a sophisticated potential model for MgO was used, which considers a number of details of the interactions within small MgO clusters. The model allows ionic shape deformations by modelling the

compressibility of individual ions as well as considering non-vanishing polarization effects in small clusters. This model is much more complex than the ones previously used in similar studies of MgO.

Free energy calculations based on this model reveal that the work of formation required for the nucleation of MgO particles to occur is at least hundreds of k_BT in the temperature range of 800 – 1200 K for an initial vapour density of 10^6 m^{-3}. Hence, MgO can be ruled out as a likely candidate for the primary nucleating dust species. It was found that meaningful nucleation rates for MgO would only be achieved in such a temperature range if the initial vapour density was much larger than some 10^{15} m^{-3}. However, such large concentrations of MgO are not found in oxygen-rich circumstellar shells.

The non-possibility of MgO nucleation is in agreement with a similar conclusion reached in the earlier studies carried out with the help of a simpler potential model [91]. Though it would be interesting to perform the free energy calculations using a novel technique based on the covariance matrix approach. Such calculations, however, were found to be very difficult and time-consuming (see Appendix B).

Chapter 7

Summary of Conclusions

The role of dust particles is well-known not only in our terrestrial atmosphere, where they are known as aerosols, but also in space where they exist around young stars and on the grander level throughout the interstellar medium. The problem of the formation and the evolutionary role of dust can be divided into three classes: dust formation, mantle growth around dust grains and the grain surface chemistry.

Heterogeneous chemistry taking place in or on the surfaces of finely divided particles within a cloud is usually described using traditional rate equations approach. In the limit where the mean number of adsorbed reactant molecules on each particle is of the order of unity, the traditional approach may lead to erroneous results. Situations of this kind are possible for ultrafine particles in the interstellar medium or in atmospheric clouds containing trace amounts of gas phase reactants. In such conditions, evolution of the the entire probability distribution of the reactant population ought to be considered, taking into account the fluctuations in the populations, instead of the evolution of the mean population.

In Chapter 3, an analytic description of such 'stochastic' chemistry was studied

through the use of master equations that govern the evolution of the population probability distribution and the results were compared with those of the traditional chemical kinetics. Two types of grain surface reactions were considered, namely $A + A \rightarrow C$ and $A + B \rightarrow C$, with the former having an analytic solution. To solve the latter problem, a number of numerical techniques were explored. It was shown that the traditional approach can overpredict a reaction rate by over one order of magnitude in typical interstellar diffuse clouds. It was found that under terrestrial atmospheric conditions, the traditional approach may well serve the purpose even though the mean reactant population may appear to fall below unity.

Heterogeneous mantle growth of the dust grains may occur when the surrounding gas molecules adsorb permanently onto the grain surface and clusters 'nucleate' heterogeneously. Again, this process of nucleation is also normally described using rate equations for the mean populations of molecular clusters. This approach is appropriate for homogeneous nucleation, but may be unsuitable in the case of heterogeneous nucleation on small particles if the mean populations are of the order of unity or less, where the statistical fluctuations in the molecular populations ought to be taken into account. A stochastic treatment of heterogeneous nucleation kinetics was presented in this thesis, in which a set of master equations were presented and a modified expression for the nucleation rate was deduced. A numerical method for solving the stochastic system was examined, with the results showing that the rate of nucleation can differ greatly from that obtained with the traditional kinetics.

The initial nucleation of the dust particles themselves in space is thought to occur in the stellar winds around stars that are undergoing heavy mass loss. One of the major candidates for the primary nucleating species was MgO because of its apparently large abundance in interstellar clouds and its high dew point. Although

earlier studies had ruled out this possibility on the grounds of its nucleation rate being too small, there was reason for sceptism about that conclusion. The potential model describing the interactions within MgO was outdated and recently, a very sophisticated potential model for MgO has become available. In Chapter 6 this potential was employed for the problem of the primary nucleating dust species in space and free energy calculations were performed in an attempt to calculate the homogeneous nucleation rate of MgO. It was found that even the better potential model gave results that rule out the possibility of MgO nucleation in circumstellar shells.

A novel technique was finally tested to perform the free energy calculations, which is meant to give results with the help of molecular dynamics simulations performed at finite temperatures. This would clearly be better than the dynamical matrix approach, which in reality yields vibrational modes of a molecule at zero temperature. This technique was, however, found to be very difficult to implement for the very small MgO clusters that were of interest in this study.

Appendix A

Properties of Modified Bessel Functions

Here we note some useful features of the modified Bessel functions appearing in Chapter 3. These can be found in any mathematical reference book such as Refs. [45, 113, 114].

If one has a second order differential equations of the form

$$r^2\frac{\mathrm{d}^2\phi}{\mathrm{d}r^2} + r\frac{\mathrm{d}\phi}{\mathrm{d}r} - (r^2 + p^2)\phi = 0, \tag{A.1}$$

the solutions to this equation are the modified Bessel functions $I_{\pm p}(r)$ and $K_p(r)$. $I_p(r)$ and $K_p(r)$ are real and positive when $p > -1$ and $r > 0$. In general, $I_p(r)$ and $I_{-p}(r)$ are linearly independent except when p is an integer. Further,

$$I_{-n}(r) = I_n(r) \quad \text{and} \quad K_{-p}(r) = K_p(r), \tag{A.2}$$

where n is an integer.

The ascending series of the function $I_p(r)$ can be written as

$$I_p(r) = \frac{(r/2)^p}{\Gamma(p+1)}\left[1 + \frac{1}{1!(p+1)}\left(\frac{r^2}{4}\right) + \frac{1}{2!(p+2)(p+1)}\left(\frac{r^2}{4}\right)^2 + \cdots\right]. \tag{A.3}$$

Here the gamma function is given by $\Gamma(p) = \int_0^\infty u^{p-1}\mathrm{e}^{-u}du$, which for integer p reduces to $\Gamma(p+1) = p!$, and is defined for $\Re p > 0$. It is also useful to note that

$$\Gamma(p+1) = p\,\Gamma(p). \tag{A.4}$$

If the argument r is very small, the Bessel function can therefore be approximated as

$$I_p(r) \sim (r/2)^p/\Gamma(p+1) \qquad \text{as } r \to 0 \quad (p \neq -1, -2, \ldots). \tag{A.5}$$

If the argument r of the Bessel function is large,

$$I_p(r) \sim \frac{\mathrm{e}^{-r}}{\sqrt{2\pi r}}\left[1 - \frac{4p^2-1}{8r} + \cdots \text{ higher order terms in } \frac{1}{r}\right] \qquad \text{as } r \to \infty. \tag{A.6}$$

For the function $K_p(r)$, the following relations hold:

$$K_p(r) \sim \frac{\Gamma(p)}{2}\frac{r}{2}^{-p} \qquad \text{as } r \to 0 \quad (\Re p > 0) \tag{A.7}$$

and

$$K_p(r) \sim \frac{\pi}{\sqrt{2r}}\,\mathrm{e}^{-r}\left[1 + \frac{4p^2-1}{8r} + \cdots \text{ higher order terms in } \frac{1}{r}\right] \qquad \text{as } r \to \infty. \tag{A.8}$$

The function $I_p(r)$ can be differentiated using the identities

$$\frac{\mathrm{d}I_p(r)}{\mathrm{d}r} = \frac{1}{2}\left[I_{p-1}(r) + I_{p+1}(r)\right] \qquad \text{and} \qquad I_{p-1}(r) - I_{p+1}(r) = \frac{2p}{r}I_p(r). \tag{A.9}$$

A similar relation for the function $K_p(r)$ has the form

$$\frac{\mathrm{d}K_p(r)}{\mathrm{d}r} = -K_{p+1}(r) + \frac{p}{r}K_p(r). \tag{A.10}$$

Appendix B

Covariance Matrix Method for Free Energy

We derived free energies of MgO clusters in Chapter 6 using the dynamical matrix approach. The molecular dynamics (MD) data produced using the CIM + PIM potential described in that Chapter can also be analysed using an alternative technique to calculate free energies of MgO clusters. This technique involves a covariance matrix method that was originally devised to calculate entropies of macromolecules [115]. The vibrational modes of a cluster are derived from the positional trajectory of the ions over a period of time. The dynamical matrix approach of §6.4 involves relaxing the crystal structure to its minimum energy state, which essentially indicates reducing the system to absolute zero temperature, and then assuming the vibrational states to be the same at a given non-zero temperature. Its applicability may therefore be questionable at high temperatures such as in stellar environments.

The covariance matrix method ought to be more suitable at high temperatures since it utilizes the ionic position history from a molecular dynamics simulation corresponding to a specific temperature. At temperatures around 300 K, this technique

has been shown to yield results with errors below 5% in physiological systems [116] and has been extensively analysed [117, 118].

Schlitter's approach assumes that each of the $3N$ degree of freedom within a cluster can be represented by a quantum mechanical harmonic oscillator. If one can obtain the positional trajectory of the ions using a molecular dynamics simulation, then the classical variance $\langle x^2 \rangle_c$ can be obtained from the history of the oscillator's Cartesian coordinate denoted by x. Here the angled brackets indicate averaging over a trajectory or over an ensemble of MD simulations. The frequency of each of the harmonic oscillator, ω, is then related to $\langle x^2 \rangle_c$ through the equipartition theorem,

$$m\omega^2 \langle x^2 \rangle_c = k_B T, \tag{B.1}$$

where m is the mass of the oscillator and T is the temperature. As a fit to the quantum mechanical properties of the harmonic oscillator, it was proposed rather heuristically that the entropy $\mathcal{S}$ of a single degree of freedom can be expressed as

$$\mathcal{S} \leq \mathcal{S}_{\mathrm{ho}} < \mathcal{S}' = \frac{1}{2} k_B \ln \left(1 + \frac{k_B T \mathrm{e}^2}{\hbar^2} m \langle x^2 \rangle_c \right), \tag{B.2}$$

where $\mathcal{S}_{\mathrm{ho}}$ is the entropy of the harmonic oscillator and $\mathrm{e} = \exp(1)$ is the Euler number. The meaning of the above formula is that the right hand side of the equation is an upper bound to the entropy of a one-dimensional quantum mechanical harmonic oscillator. According to Equation (B.2), $\mathcal{S}'$ approaches zero when the temperature is reduced to zero and is proportional to $\ln T$ for $T \to \infty$. Both of these features satisfy the requirements of the quantum mechanical and the classical limits respectively.

For many degrees of freedom, the formula is generalised using the *covariance matrix* $\boldsymbol{\sigma}$ instead of the variance $\langle x^2 \rangle_c$, so that

$$\mathcal{S} < \mathcal{S}' = \frac{1}{2} k_B \ln \det \left[\mathbf{1} + \frac{k_B T \mathrm{e}^2}{\hbar^2} \mathbf{M}^{1/2} \boldsymbol{\sigma} \mathbf{M}^{1/2} \right], \tag{B.3}$$

where the elements of the matrix $\boldsymbol{\sigma}$ are given by

$$\sigma_{\ell j} = \langle\, (x_\ell - \langle x_\ell \rangle)(x_j - \langle x_j \rangle)\,\rangle. \tag{B.4}$$

Here $x_1, x_2, \ldots, x_{3N}$ are the Cartesian coordinates of the N ions in the system.

Using a molecular dynamics simulation, an ensemble of ionic positions can be obtained over a period of time to determine $\boldsymbol{\sigma}$. The matrix $\mathbf{M}^{1/2}\boldsymbol{\sigma}\mathbf{M}^{1/2}$ is then symmetric and semidefinite, so that its eigenvalues can be obtained through diagonalization. These eigenvalues, λ_k^{cov} are the classical variance $\langle q_j^2 \rangle_c$ of the resulting new, uncorrelated coordinates q_j. Contrasting this with the one-dimensional case, $m\langle x^2 \rangle_c$ is now replaced by λ_k^{cov}, so using Equations (B.1) and (4.26), it can be shown that the vibrational free energy from the covariance matrix method is expressed as

$$\mathcal{F}_{\text{vib}}^{\text{cov}} \;=\; -\frac{1}{2} k_B T \ln\left(\prod_{k=1}^{6i-6} \frac{k_B T}{\hbar^2} \lambda_k^{\text{cov}} \right). \tag{B.5}$$

One can use this expression instead of the one given in Equation (4.26) and perform the rest of the calculations as in Chapter 4 to obtain the work of formation of an $i-$ cluster. This was attempted and found to be very difficult due to mainly two reasons. Due to the fact that in astrophysical conditions the size of the critical cluster is of the order of 10, or maybe even smaller [85], the calculations needed to be carried out in the small size regime. This makes it difficult to accurately define the temperature of the cluster in an MD simulation and leads to a rather unreliable covariance matrix. Also, the clusters had to be simulated at the high temperatures of circumstellar clouds, so in the MD simulation the total energy of the system would settle down to a stable value after quite long time. The combined effect of these two difficulties was that in order to have reliable results, simulation for each temperature would have required an enormously long execution time with the current generation of computers and was therefore not pursued further.

Bibliography

[1] E. L. O. Bakes, *The Astrochemical Evolution of the Interstellar Medium*, Twin Press, Vledder, 1997.

[2] M. A. Dopita and R. S. Sutherland, *Astrophysics of the Diffuse Universe*, Springer, New York, 2003.

[3] M. Zelik, A. Gregory, and E. v. P. Smith, *Introductory Astronomy and Astrophysics*, Saunders HBJ, Orlando, 3rd ed., 1992.

[4] J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics*, Wiley, New York, 1998.

[5] A. A. Lushnikov, J. S. Bhatt, and I. J. Ford, *J. Aerosol. Sci.*, **34**, 1117 (2003).

[6] J. S. Bhatt and I. J. Ford, *J. Chem. Phys.*, **118**, 3166 (2003).

[7] P. Coles and F. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*, Wiley, Chichester, 1995.

[8] A. J. Meadows, *Stellar Evolution*, Pergamon, Oxford, 1978.

[9] M. Inglis, *Observer's Guide to Stellar Evolution*, Springer, London, 2003.

[10] E. Krügel, *The Physics of Interstellar Dust*, IOP, London, 2003.

[11] W. S. C. Williams, *Nuclear and Particle Physics*, Oxford University Press, Oxford, 1991.

[12] L. H. Aller, *Atoms, Stars and Nebulae*, Cambridge University Press, Cambridge, 3rd ed., 1991.

[13] B. E. J. Pagel, *Nucleosynthesis and Chemical Evolution of Galaxies*, Cambridge University Press, Cambridge, 1997.

[14] G. Wynn-Williams, *The Fullness of Space*, Cambridge University Press, Cambridge, 1992.

[15] D. C. B. Whittet, *Dust in the Galactic Environment*, IOP, London, 2nd ed., 2003.

[16] R. Bowers and T. Deeming, *Astrophysics II: Interstellar Matter and Galaxies*, Jones and Bartlett, London, 1984.

[17] N. Kaifu, *Molecular Processes in Space*, Plenum, New York, 1990.

[18] D. A. Williams, *Faraday Discuss.*, **109**, 1 (1998).

[19] D. Flower, *Molecular Collisions in the Interstellar Medium*, Cambridge University Press, Cambridge, 1990.

[20] R. J. Gould and E. E. Salpeter, *Astrophys. J.*, **138**, 393 (1963).

[21] A. J. Farebrother, A. J. H. M. Meijer, D. C. Clary, and A. J. Fisher, *Chem. Phys. Lett.*, **319**, 303 (2000).

[22] M. Allen and G. W. Robinson, *Astrophys. J.*, **212**, 396 (1977).

[23] A. G. G. M. Tielens and W. Hagen, *Astron. Astrophys.*, **114**, 245 (1982).

[24] A. G. G. M. Tielens and L. J. Allamandola In *Interstellar Processes*, D. J. Hollenbach and H. A. Thornson, Eds.; Reidel, Dordrecht, 1987.

[25] S. B. Charnley, A. G. G. M. Tielens, and S. D. Ridgers, *Astrophys. J.*, **482**, L203 (1997).

[26] A. G. G. M. Tielens and S. B. Charnley, *Origins of Life and Evolution of the Biosphere*, **27**, 34 (1997).

[27] J. B. Pickles and D. A. Williams, *Astrophys. Space Sci.*, **52**, 443 (1977).

[28] L. B. d'Hendecourt, L. J. Allamandola, and J. M. Greenberg, *Astron. Astrophys.*, **152**, 130 (1985).

[29] P. D. Brown and S. B. Charnley, *Mon. Not. R. Astron. Soc.*, **244**, 432 (1990).

[30] T. I. Hasegawa, E. Herbst, and C. M. Leung, *Astrophys. J. Suppl.*, **82**, 167 (1992).

[31] K. Willacy and D. A. Williams, *Mon. Not. R. Astron. Soc.*, **260**, 635 (1993).

[32] O. M. Shalabiea and J. M. Greenberg, *Astron. Astrophys.*, **290**, 266 (1994).

[33] D. T. Gillespie, *J. Comp. Phys.*, **22**, 403 (1976).

[34] D. T. Gillespie, *J. Phys. Chem.*, **81**, 2340 (1977).

[35] M. D. Hatlee and J. J. Kozak, *J. Chem. Phys.*, **72**, 4358 (1980).

[36] M. D. Hatlee and J. J. Kozak, *J. Chem. Phys.*, **74**, 1098 (1981).

[37] M. D. Hatlee and J. J. Kozak, *J. Chem. Phys.*, **74**, 5627 (1981).

[38] M. Mozurkewich, *Geophys. Res. Lett.*, **24**, 3209 (1997).

[39] S. B. Charnley, *Astrophys. J.*, **509**, L121 (1998).

[40] P. Caselli, T. I. Hasegawa, and E. Herbst, *Astrophys. J.*, **495**, 309 (1998).

[41] O. M. Shalabiea, P. Caselli, and E. Herbst, *Astrophys. J.*, **502**, 652 (1998).

[42] D. P. Ruffle and E. Herbst, *Mon. Not. R. Astron. Soc.*, **319**, 837 (2000).

[43] A. A. Lushnikov, *J. Aerosol Sci.*, **30: Suppl. 1**, S607 (1999).

[44] N. J. B. Green, T. Toniazzo, M. J. Pilling, D. P. Ruffle, N. Bell, and T. W. Hartquist, *Astron. Astrophys.*, **375**, 1111 (2001).

[45] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, 5th ed., 1994.

[46] W. H. e. a. Press, *Numerical Recipes in Fortran*, 2nd ed.

[47] J. F. Pankow, *Atmospheric Environment*, **21**, 2275 (1987).

[48] J. F. Pankow and T. F. Bidleman, *Atmospheric Environment*, **26A**, 1071 (1992).

[49] I. A. Crawford, private communication.

[50] D. Kashchiev, *Nucleation: Basic Theory with Applications*, Butterworth-Heinemann, Oxford, 2000.

[51] D. T. Wu, *Solid State Physics - Advances in Research and Applications*, **50**, 37 (1997).

[52] J. J. Burton In *Statistical Mechanics*, B. J. Berne, Ed.; Plenum, New York, 1977.

[53] I. J. Ford, *Proc. Instn. Mech. Engrs.*, **218**(C), 883 (2004).

[54] J. V. Narlikar and T. Padmanabhan, *Ann. Rev. Astron. Astrophys.*, **29**, 325 (1991).

[55] Y. B. Zel'dovich, *Mon. Not. R. Astron. Soc.*, **160**, 1P (1972).

[56] R. Becker and W. Döring, *Ann. Phyz., Leipzig*, **24**, 719 (1935).

[57] I. J. Ford, *Phys. Rev. E*, **56**, 5615 (1997).

[58] J. R. Hook and H. E. Hall, *Solid State Physics Physics*, Wiley, Chichester, 2nd ed., 1991.

[59] F. Mandl, *Statistical Physics*, Wiley, London, 2nd ed., 1988.

[60] F. F. Abraham, *Homogeneous Nucleation Theory*, Academic Press, New York, 1974.

[61] M. J. E. and M. G. Mayer, *Statistical Mechanics*, Wiley, New York, 2nd ed., 1977.

[62] N. H. Fletcher, *J. Chem. Phys.*, **29**, 572 (1958).

[63] N. H. Fletcher, *J. Meteor.*, **16**, 173 (1959).

[64] N. H. Fletcher, *J. Chem. Phys.*, **31**, 1136 (1959).

[65] N. H. Fletcher, *Phil. Mag.*, **7**, 255 (1962).

[66] H. R. Pruppacher and J. D. Klett, *Microphysics of Clouds and Precipitation*, Kluwer Academic, Dordrecht, 1997.

[67] B. Gorbunov and R. Hamilton, *J. Aerosol Sci.*, **28**, 239 (1997).

[68] B. Gorbunov, *J. Chem. Phys.*, **110**, 10035 (1999).

[69] W. Ebeling, L. Schimansky-Geier, and F. Schweitzer, *Z. Phys. Chem.*, **169**, 1 (1990).

[70] S. Manjunath, K. S. Gandhi, R. Kumar, and D. Ramkrishna, *Chem. Eng. Sci.*, **49**, 1451 (1994).

[71] J. E. Dyson and D. A. Williams, *The Physics of the Interstellar Medium*, IOP, London, 1997.

[72] F. Hoyle and N. C. Wickramasinghe, *Mon. Not. R. Astron. Soc.*, **124**, 417 (1962).

[73] M. F. Bode In M. E. Bailey and D. A. Williams, Eds., *Dust in the Universe*, Cambridge, 1988. Cambridge University Press.

[74] H.-P. Gail and E. Sedlmayr In *Physical Processes in Interstellar Clouds*, G. E. Morfill and M. Scholer, Eds.; D. Reidel, Dordrecht, 1987.

[75] A. K. Dupree, *Annu. Rev. Astron. Astrophys.*, **24**, 377 (1986).

[76] M. Jura and S. G. Kleinmann, *Astrophys. J. Suppl.*, **73**, 769 (1990).

[77] E. Sedlmayr and C. Dominik, *Space Science Reviews*, **73**, 211 (1995).

[78] G. R. Knapp and M. Morris, *Astrophys. J.*, **292**, 640 (1985).

[79] H. A. Thronson, J. H. Black, J. Bally, and P. Hacking, *Astrophys. J.*, **322**, 770 (1987).

[80] H.-P. Gail and E. Sedlmayr, *Astron. Astrophys.*, **148**, 183 (1985).

[81] H.-P. Gail and E. Sedlmayr, *Astron. Astrophys.*, **206**, 153 (1988).

[82] T. Tsuji, *Astron. Astrophys.*, **23**, 411 (1973).

[83] J. A. Nuth and J. H. Hecht, *Astrophys. Space Sci.*, **163**, 79 (1990).

[84] B. Begemann, T. Henning, H. Mutschke, and J. Dorschner, *Planet. Space Sci.*, **43**, 1257 (1995).

[85] A. B. C. Patzer, T. M. Köhler, and E. Sedlmayr, *Planet. Space Sci.*, **43**, 1233 (1995).

[86] H.-P. Gail and E. Sedlmayr, *Astron. Astrophys.*, **166**, 225 (1986).

[87] H.-P. Gail and E. Sedlmayr, *Faraday Discuss.*, **109**, 303 (1998).

[88] J. A. Nuth and B. Donn, *J. Chem. Phys.*, **77**, 2639 (1982).

[89] J. A. Nuth and B. Donn, *J. Chem. Phys.*, **78**, 1618 (1983).

[90] S. MacLean and W. W. Duley, *Astrophys. J.*, **252**, L25 (1982).

[91] T. M. Köhler, H.-P. Gail, and E. Sedlmayr, *Astron. Astrophys.*, **320**, 553 (1997).

[92] M. Wilson, N. C. Pyper, and J. H. Harding, *J. Chem. Phys.*, **104**, 8068 (1996).

[93] M. Wilson and P. A. Madden, *J. Phys. Chem.*, **100**, 1227 (1996).

[94] M. Wilson, *J. Phys. Chem. B*, **101**, 4917 (1997).

[95] A. J. Rowley, P. Jemmer, M. Wilson, and P. A. Madden, *J. Chem. Phys.*, **108**, 10209 (1998).

[96] M. Born and J. E. Mayer, *Zeitschr. f. Phys.*, **75**, 1 (1932).

[97] E. S. Rittner, *J. Chem. Phys.*, **19**, 1030 (1951).

[98] C. T. O'Konski, *J. Chem. Phys.*, **23**, 1174 (1955).

[99] P. Brumer and M. Karplus, *J. Chem. Phys.*, **58**, 3903 (1973).

[100] L. Pauling, Cornell University Press, 3rd ed., 1960.

[101] P. J. Ziemann and A. W. Castleman, *J. Chem. Phys.*, **94**, 718 (1991).

[102] B. G. Dick and A. W. Overhauser, *Phys. Rev.*, **112**, 90 (1958).

[103] U. Schröder, *Solid State Commun.*, **4**, 347 (1966).

[104] L. L. Boyer, M. J. Mehl, J. L. Feldman, J. R. Hardy, J. W. Flocken, and Y. Fong, *Phys. Rev. Lett.*, **54**, 1940 (1985).

[105] N. C. Pyper, *Phil. Trans. R. Soc. London A*, **320**, 107 (1986).

[106] K. T. Tang and J. P. Toennies, *J. Chem. Phys.*, **80**, 3726 (1984).

[107] J. H. Harding and N. C. Pyper, *Philos. Mag. Lett.*, **71**, 113 (1995).

[108] M. Wilson and P. A. Madden, *J. Phys.: Condens. Matter*, **5**, 6833 (1993).

[109] M. Wilson and P. A. Madden, *J. Phys.: Condens. Matter*, **6**, 159 (1994).

[110] M. Wilson, P. A. Madden, and B. J. Costa-Cabral, *J. Phys. Chem*, **100**, 1227 (1996).

[111] A. D. Buckingham, *Adv. Chem. Phys.*, **12**, 107 (1967).

[112] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders, Philadelphia, 1976.

[113] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*, Dover, New York, 1972.

[114] D. H. Menzel, Ed., *Fundamental Formulas of Physics*, Vol. 1, Dover, New York, 1960.

[115] J. Schlitter, *Chem. Phys. Lett.*, **215**, 617 (1993).

[116] H. Schäfer, A. E. Mark, and W. F. v. Gunsteren, *J. Chem. Phys.*, **113**, 7809 (2000).

[117] H. Schäfer, X. Daura, A. E. Mark, and W. F. v. Gunsteren, *Proteins: Structure, Function, and Genetics*, **43**, 45 (2001).

[118] I. Andricioaei and M. Karplus, *J. Chem. Phys.*, **115**, 6289 (2001).